

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 14:43:51 ; Search time 85.35 Seconds
(without alignments)
4990.375 Million cell updates/sec

Title: US-09-710-339-1

Perfect score: 1734
Sequence: 1 tcaatcagcttcctccctc.....aaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCRTUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfilst1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	375.4	21.6	1404	US-08-204-656B-7	Sequence 7, Appl
2	375.4	21.6	1404	US-08-470-702-4	Sequence 4, Appl
3	375.4	21.6	1404	US-08-467-831-4	Sequence 4, Appl
4	373.8	21.6	1404	US-08-204-656B-1	Sequence 1, Appl
5	373.8	21.6	1404	US-08-204-656B-3	Sequence 3, Appl
6	373.8	21.6	1404	US-08-204-656B-5	Sequence 5, Appl
7	373.8	21.6	1404	US-08-470-702-1	Sequence 1, Appl
8	373.8	21.6	1404	US-08-470-702-2	Sequence 2, Appl
9	373.8	21.6	1404	US-08-470-702-3	Sequence 3, Appl
10	373.8	21.6	1404	US-08-467-831-1	Sequence 1, Appl
11	373.8	21.6	1404	US-08-467-831-2	Sequence 2, Appl
12	373.8	21.6	1404	US-08-467-831-3	Sequence 3, Appl
13	68.6	4.0	93	US-09-189-060B-63	Sequence 63, Appl
14	68.4	3.9	919	US-09-189-060B-67	Sequence 67, Appl
15	68.4	3.9	922	US-09-189-060B-65	Sequence 65, Appl
16	68.4	3.9	922	US-09-189-060B-73	Sequence 73, Appl
17	68.4	3.9	924	US-09-189-060B-69	Sequence 69, Appl
18	63.4	3.7	912	US-09-189-060B-71	Sequence 71, Appl
19	58	3.3	2582	US-08-816-105A-2	Sequence 2, Appl
20	55.4	3.3	5163	US-08-700-651-1	Sequence 1, Appl
21	55.4	3.2	5163	US-08-928-361B-4	Sequence 4, Appl
22	55.4	3.2	5318	US-08-700-651-2	Sequence 2, Appl
23	55.4	3.2	5318	US-08-928-361B-3	Sequence 3, Appl
24	51.4	3.0	5511	US-08-928-361B-2	Sequence 2, Appl
25	51.4	3.0	7334	US-08-928-361B-1	Sequence 1, Appl
26	50.6	2.9	2160	US-09-386-607-1	Sequence 1, Appl
27	49.2	2.8	1295	US-08-433-854-3	Sequence 3, Appl

28	49.2	2.8	1295	US-08-174-745A-3	Sequence 3, Appl
29	49.2	2.8	1295	US-08-195-947-3	Sequence 3, Appl
30	49.2	2.8	1295	US-08-433-885-3	Sequence 3, Appl
31	49.2	2.8	1295	US-08-433-908B-3	Sequence 3, Appl
32	49.2	2.8	1295	US-08-410-614-3	Sequence 3, Appl
33	48.4	2.8	390	US-09-197-649-7	Sequence 7, Appl
34	44.8	2.6	1100	US-07-861-458C-4	Sequence 4, Appl
35	44.8	2.6	2781	US-08-140-008A-1	Sequence 1, Appl
36	43.8	2.5	2061	US-08-204-656B-9	Sequence 9, Appl
37	43.8	2.5	2061	US-08-470-702-5	Sequence 5, Appl
38	43.8	2.5	2061	US-08-467-831-5	Sequence 5, Appl
39	43.2	2.5	5580	US-08-814-052-10	Sequence 10, Appl
40	43.2	2.5	5580	US-08-812-829-10	Sequence 10, Appl
41	43.2	2.5	5679	US-08-814-052-9	Sequence 9, Appl
42	43.2	2.5	5679	US-08-812-829-9	Sequence 9, Appl
43	43.2	2.5	5697	US-08-814-052-11	Sequence 11, Appl
44	43.2	2.5	5697	US-08-812-829-11	Sequence 11, Appl
45	43	2.5	2209	US-08-514-014-1	Sequence 1, Appl

ALIGNMENTS

```
RESULT 1
: US-08-204-656B-7
: Sequence 7, Application US/08204656B
: Patent No. 5538882
:
: GENERAL INFORMATION:
:
: APPLICANT: Matsui, Ikuo
: APPLICANT: Ishikawa, Kazuhiko
: APPLICANT: Miyairi, Sachio
: APPLICANT: Honda, Koichi
:
: TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
: TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
: TITLE OF INVENTION: Oligosaccharide Using The Enzyme
:
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
: STREET: 8110 Gatehouse Road, Suite 500 East
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22042
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/204,656B
: FILING DATE: 02-MAR-1994
: CLASSIFICATION: A35
: ATTORNEY/AGENT INFORMATION:
: NAME: Weiner, Marc S.
: REGISTRATION NUMBER: 32,181
: REFERENCE/DOCKET NUMBER: 234-252P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 205-8000
: TELEFAX: (703) 205-8050
:
: TELEEX: 248345
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1404 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: other nucleic acid
: DESCRIPTION: /desc = "Synthetic DNA"
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: IMMEDIATE SOURCE:
: CLONE: Derived from plasmid pSF11 (Agric. Biol. Chem.
: FEATURE:
```

```

; NAME/KEY: CDS
; LOCATION: 1..1404 /note= "Nucleotides 1-1404
; OTHER INFORMATION: correspond to nucleotides 79-1482 of the Saccharomyces
; OTHER INFORMATION: fibuligera `amylase structural gene"
; US-08-204-656B-7

```

```

Query Match          21.6%; Score 375.4; DB 1; Length 1404;
Best Local Similarity 55.6%; Pred. No. 5,6e-98;
Matches 752; Conservative 0; Mismatches 586; Indels 15; Gaps 1;

```

```

OY 121 ctgcgagctggcgatcgcatccatcttctcctcgaagctgattgcaagagcg 180
Db 11 CTGATTAATGAGGATCAGACGCTATTTATCAAAATGTCACGACGATTTCTAGAACG 70
OY 181 atggctgagcagctgagctgtaattactgagatcagaatctggtggtgaacatg 240
Db 71 ATGGGATATACAGATGCTCTGTAACACAGAAAGATGACCTTTACTGTGTGTTCTTCC 130
OY 241 agggatcatcgacagatgagctactatccagggatgagctgcaagccatcgatca 300
Db 131 AAGGATCATTAAGAAAGTTGGATTACATCAAGATATGGGCTTACTGCTATTTGATTT 190
OY 301 ccccgcttaacagccagctgcccagaccagcatatgagatgagctacatgctact 360
Db 191 CTCCAGTTGTGAAAAACATTCCTCCGATACACAGCATATGTTATGCTTATCATGTA 250
OY 361 ggcacagatatctactctctgaacgaacatcgagcagctgagatgagctgaagcgc 420
Db 251 GGATGAAACATATACAAATTAATGAACCTTTGGTACTGCTGATGATTTGAAGTCTT 310
OY 421 tctctcgcccttcataagagggagtgatctctatgctgagtggttgctacata 480
Db 311 TGGCACAAGAAATGCGATGATGATGTTAATGTTGATGATGATGATGATGATGAT 370
OY 481 tgggtatgagtgagggagtgatccagctgagctgagtggttcaaacgctgagctcc 540
Db 371 ACGGAGATGAGGAGTGGAGATGATGATGATGATGATGATGATGATGATGATGAT 430
OY 541 aagactactccacccgctgctgcttcatcaaaactatgaagatcagatcagtgag 600
Db 431 AAAATACTTCCATATACACTACTGCTTATTTCAAAATGATGATGATGATGATGAT 490
OY 601 attgctgctagagataacactgctcctgctgctgctgctgctgctgctgctgct 660
Db 491 GTTGCTGGGAGGAGTCACTCTTCACTTGCATTACAGATTGAGAAACGAAATAGCAG 550
OY 661 tcaagaatgagtgagcagctggttgagatcatggtatcgactactccatcgagcc 720
Db 551 TGGCTCTCAGTTTTCATTTCTTGGGTTAAAGATTTTGTGGCAATTTACTCATTT 610
OY 721 tccglatcgacacagtaaacacgctccagaagagctctgcccgggttacacaacg 780
Db 611 TAAGATATGATAGTGTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 670
OY 781 cagcgctgactgctgagcgaggtgctgagcggtgctgagcgctgactgctgctacc 840
Db 671 CTGGAGTTTACTCAGTAGGCGAAGTTTCCAAAGGAGACCCAGCTTATACATGCCC 730
OY 841 agaagctgagcagcggtgactgagctactccatcttacttaccactctcctcaagc 900
Db 731 AAAATTTACATTTCCAGGGGTTAGTAAATTTATTCATTTGTTACCCAAACGAGAT 790
OY 901 agtaacctccgagcagatgagcagctctacacacgcatcaaacgctcaaatcgact 960
Db 791 AAATCTAGTATCAAGTTCCAGTGATGATGATGATGATGATGATGATGATGATGAT 850
OY 961 gtccagactcaaacactcctgggacatcgctcgaggaacacagacaacgagtgctc 1020
Db 851 GTTGAGTCCAACTTTGTTGACAAACTTTGAGAAATTCAGATATGAAAGATTCGCT 910
OY 1021 cttaacacaagacatagcctctgccaagacgctcgacagatctatctcctcaagc 1080

```

```

Db 911 CAATGACACAGCCACCAAGTTGATTTCTAATGCTATTCATTTGCTTTGGGGGATG 970
OY 1081 gaatcccatcatcatcgcgcgcgaagaacagactacgctggcggaacgcccgcga 1140
Db 971 GATTCCTGCTATTTACTATGACCAAGAAACAGGCTTGAGCGAAACAGTACCAACA 1030
OY 1141 accggaagcaacctgctgctcgtgagctaccgcagcagcagcagcagctgtaacgt 1200
Db 1031 ACAGAGAGGCTGTTGTGTTATCCGGCTACACAAAGAGTGATCTATTAACAGCTC 1090
OY 1201 cctcgcgaacgcaacccggaactatgcatatgacaagaatagatgctgtaacct 1260
Db 1091 CCAAGCTAATGCTGCCAGAAAGCCGCTTATATCAAGATCAAGCTATGACCTGCG 1150
OY 1261 agaacgccccatctacaaagacacacagatgcgcaltgccaagcgcaagatgg 1320
Db 1151 AGCTTCTGTGATCTTTTCAATGACCATGTTATTTCAACAAAGAGGACGCTGTT 1210
OY 1321 cgcagatcgtactatctgtccacaaggggtgctgagtgatgctatccctcct 1380
Db 1211 CTG-----TTTCAACACCTTGCTTCCAGCGGTTCTTCTGATGATGACTA 1255
OY 1381 tgaagtgtcggttacaacagccgcaagcaattgaacgaggtcattgctgacagc 1440
Db 1256 TTTCCAAACACAGGCTTACGTTCCGGTGAGGATTTGGTAGAAGTTTGATGATG 1315
OY 1441 tgaagtggtgagtggaatgctgctc 1473
Db 1316 TTAGCGGACGCTGACTTACAAAGTTTCTATCC 1348

```

```

RESULT 2
US-08-470-702-4
; Sequence 4, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Catehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,702
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,656
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 4:

```



```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; HYPOTHEITICAL: NO
; ANTI-SENSE: NO
; US-08-470-702-4

```

```

Query Match      21.6%; Score 375.4; DB 1; Length 1404;
Best Local Similarity 55.6%; Pred. No. 5.6e-98;
Matches 752; Conservative 0; Mismatches 586; Indels 15; Gaps 1;

```

```

QY 121 ctggcgactgagctgcgaatccatcttctccctccgaagatcggattggcaagcg 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 11 CTGTAATGAGATGACAGCTATTTATCAAAATGTCACGACAGATTGGCTAGAACCG 70
QY 181 atggatcgacgactgcgaactgtlaatactgcgagatcagaataactgtgtggaacatg 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 71 ATGGGATACAGAGGCTTCCGTAAACAGAGATAGACTTACTGTGGGTCTTCC 130
QY 241 agggcatatcagacaagltgactatatccaggaatgaggcttcacagccatctgtatca 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131 AAGCATCATTAAGAGATGGATTACATCAAAAGATATGGGCTTACTGCTATTGGATT 190
QY 301 ccccgatcaagcccaagctgcccagacccgcatatgagagatgctacacatgact 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 CTCACGTTGTGAAACATTCCTCCGATTAACACGATATGGTTATGCTTATCATGTA 250
QY 361 ggcagcagatatatactctctgaacgaaactacgagcactgcagactgagactggaagcgc 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 251 GGATGAAGAACAATATACAAATTAATGAACCTTGGTACGCTGATGATTTGAAGCTT 310
QY 421 tctctcgccctcattgagagggggtatctatgtgtcgatgtgtgttaacata 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 311 TGGCACAAGATTGACAGATGATATGTTAATGGTGGATATGCTTACCAACATT 370
QY 481 tgggctatgagggcggtgagccagactgagatcaggtgtttaaaccgttcctccc 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 371 ACGGACGATGAGGAGTGAGATATGATTCGATTCAGACTACACCCCGTTCAACAGCC 430
QY 541 aagactatccaccgcttcgttccatcaaaactaagaatcagactcagactcaggttgaag 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 431 AAAAGTACTTCATTAAGTACTGTCTTATTTCAAACTATGATGACCAAGCTCAGGTTCAA 490
QY 601 attgctgctagggagataaacctgtctccttgctgactcgcataccaaccaagagtgg 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 491 GTTGTGGGAAGGTGACTCTTCAGTTGATTCAGATTTGAGAACGGAAGATAGCGACG 550
QY 661 tcaagaagaatgtagaagcagctgggtgagatcatgtagatcgaactcaatcagagcc 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 551 TGGCTCAGTTTTCATTTCTTGGGTTAAAGATTTTGTTGGCAATTTACTCAATATATG 610
QY 721 tccgatalcagacaglaaacaacgltccagaagactctcgtgcccgggtacaaacagccg 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 611 TAAGATATGATAGTCTTAACATGTGACCAAGGCTTTTCCCGATTTTGTATGATGAT 670
QY 781 caggcggtactgtatcggcgaggtgctcgagagtgatccggcgtctacactgttccctacc 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 671 CTGGAGTTTACTCACTAGCGCAAGCTTTTCAAGGAGACCCAGCTTATATACAGCCAT 730
QY 841 agaagcatalcagagcggtactgaactatcccatctatccactcctcgaagccctca 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 AAAATTAATTCAGAGGGTTAGTATTTATTCATTTACTACCAACAGCAGATTTT 790
QY 901 agtcaacctccggcagcatgtagcagcctctcaacatgatcaacaacggtcaaatccgact 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 791 AAATCTAGTATTCAGATTCAGATGAGTACCTCAAAATGATTTTCAAGCGTCTCTCA 850
QY 961 gtccagactcaaacactcctggtgacatctgcgcgagaacaacagacccaacccagttcg 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 851 GTTGGATCAACCTTTGTTGACAAACTTTGTAGAAAATCAGATATGAAGGTTGCTT 910
QY 1021 cttaaccaagacatagccctcgccaagaacgtgcgaacatctcatcctcaacagc 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 911 CAATGACAGCAGCAAAATTTGATTTCTAATGCTATTCATTTGCTTTGGGTGATG 970
QY 1081 gaatcccatctatcagccggcccaagaacagcactacgagcgcggaacgaaccccgga 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 971 GATTCCTGCTATTTACTATGAGCAAGAACAGGCTTGAGCGGAAAAGTGACCAACA 1030
QY 1141 accggaagcaactggtcctcgtggtctaccgagccgagcgagctgtatcaagttaattg 1200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1031 ACAGAGAGGCTTGTGTTATCCGCTTACACAAAGAGAGTACTATTATCAAGCTCAT 1090
QY 1201 cctccggaagcaactcgaactatgcatatagcaagatacagatcgtgacctca 1260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1091 CCAAGCTATATGCTGCAAGAAAGCCGCGCTTATCAAGACTCAAGCTATGACACTGCG 1150
QY 1261 agaatgccccatctacaagaacgacaacagatcgccatgcccgaagggcacagattgg 1320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1151 AGCTTCTGTGATCTTTCAAAATGACCATGTTATGCAACAAAAGAGCAGCGTTGTT 1210
QY 1321 cgcagatcgtgactatctgttcaacaaggtgcttcggtgattcgtataccctcct 1380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1211 CTG-----TTTTCACACACCTTGTTCCAGCGGTTCTTCTGTGATGTGACTA 1255
QY 1381 tgaatggtgaggttacaacagccgagcaaatgtagaggtcattgctgcagacgc 1440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1256 TTTCCAAACACAGGTTACAGTCCGCTGAGAGATTGGTAAAGTTTGGACATGACGTAC 1315
QY 1441 tgaaggtggttcgagatggaatgagctgtc 1473
Db 1316 TTAGCGGACAGCTGCTGACTTACAGTTTCTATCC 1348

RESULT 3
US-08-467-831-4
; Sequence 4, Application US/08467831
; Patent No. 5635378
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Galenhouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,831
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,656
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000

```



```

STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: Derived from plasmid pSI\1 (Agric. Biol. Chem.
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1404
OTHER INFORMATION: /note= "Nucleotides 1-1404
correspond to nucleotides 79-1482 of the Saccharomyces
OTHER INFORMATION: fibuligera `'-amylase structural gene"
US-08-204-656B-5

Query Match      21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 1.6e-97;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY 121 ctgcgagctggcgagtcgcaatcattatctcctctcagcgatgattgcaaggagcg 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11 CTGATTAATGAGATGACACAGTCTATTATCAAAATGCTCACTGACAGATTGCTAGAACG 70

QY 181 atgggtcgcagactgcgactgtaactacgcatcagaataactctgtggaacatgagc 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 71 ATGGTGATTAAGAGTCTCTCTGTACACACAGATAGACTTTACTGTGGTCTTTCTTCC 130

QY 241 agggcatcaccgaagcttgactatataccaggaaatgggcttcacagccatcgtatca 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 131 AAGGCATCAATAAGAGTGTGATTACATCAAGATATGGCTTTACTGCTATTGGATT 190

QY 301 ccccgcttacagccagctgccccagacacgcacatgtagagatgctaccatgagctact 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 191 CTCACAGTTGTTGAAAACATTCGCCGATTAACACACATATGTTTGTCTATCAATGCTCT 250

QY 361 ggcaagagatatatactctctgaacaaactacgacgacatcagatgacttgaagcgcg 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 251 GGATGAAGAACATTTACAAAATTAATGAATAAATTTGGTACTGCTGATGATTTGAAGCT 310

QY 421 tctctcgagccctcatgagaggggagatglatcttatgctcgatggtgttgctaacata 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 311 TGGACAAAGAAATTCACAGATCGTGTATGTGTTAATGTGATATCGTACCAACCAT 370

QY 481 tggactatgtagggagggtagcaccatgcatatcacagtgtgtaaacccgttaacgtccc 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 371 ACGGACAGTAGGCAAGTGGAGATGATGATCTGATTACTCAGAGTACACCCCGTTACAGACC 430

```

```

QY 541 aagactactccaccgcctctcttcatcaaaaactatgaatgaatcagctcagttgag 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 431 AAAAGTACTTCCATTAATCTACTGTCTATTCAAACTATGATGATCCCAAGCTCAGTTCAA 490

QY 601 attcctgcttaggagataaactgtctcctctgctcgatcgtatcaccacgaagtgtg 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 491 GTTGCTGGGAAGGTGACTCTCTGATGCAATTCACAAATTTGGAAGAGGAAGATTAACGACG 550

QY 661 tcaagaatgatgtgctgactgagctgggtggatcattgttatcgacactcattcagcgcc 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 551 TGCCCTCAGTTTCAATTTCTTGGGTTAAAGATTTTGTGGCAATTTACTCAATTTGATGTG 610

QY 721 tccgatalcgacacagtaaaacacgctccagaaggactctgccccgggtacaacaaagcg 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 611 TAAGAAATGATAGTGGCAAAACATGTGGACCAAGCTTTTCCCGATTTTGTAGTGCAT 670

QY 781 caggcgtgactgtatcgcgagagtgctcgaagtgatccggcctcacttgctccctacc 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 671 CTGGAGTTTACTCAGTATGCGGAAGTTTCCAAAGGAACCCAGCTTATACATGCGCATACC 730

QY 841 agaacgctcagagcggtgactgaactatccattactatccactcctcaacgcttca 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 731 AAAATTTCAATTTCCAGGGGTAGTATATTCATTTGATCTCCACACAGAGATTTT 790

QY 901 agtcaacctccggcagcatgtagcagacctctacaacatgatcaaacocgltcaatccgact 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 791 AAATCTCATGATTCAAGTTCCAGTGAAGTTGACTCAAAATGATTCAAGGCTTGCCTACGAT 850

QY 961 gtccagactcaaacatctctggtgacatctgtcgagaacccacgaaccccggttcgctt 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 851 GTTCGATTCACAACTTTGTGCAAACTTTGTAGAAAATTCAGATTAAGAAAGGTGCGTT 910

QY 1021 cttaacccaagagatagccctcgccaagaacgctgcagacatcatcatcctcaacagcg 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 911 CAATGACCAAGGACCAAAAGTTGATTTCTATGCTATTCATTTGCTTTGGGTGATG 970

QY 1081 gaatcccatcatcctacgcccggccaagaagaacacactcgcggcggaagaccccgga 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 971 GTATTCTCGTATTACTATGATGACAAAGACAAAGGCTTGAGCGGAAAAAGTGACCAAAACA 1030

QY 1141 accgcgaagaaactgtgtctccgggttaaccgacggaagagctgtacaagttaattg 1200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1031 ACAGAGAGGCGCTTGTTGGTTATTCGCTCAACAAAGAGAGATATTAACAAGCTCAATG 1090

QY 1201 cctcgcgaacgcaatccgcgaatgcatgaagaagatcacagatcgttgactctaca 1260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1091 CCAAGGCTAATGCTGCCAGAAAGCGCGCTTTTATCAAGACTCAAGCTATGCTACACTCGC 1150

QY 1261 agaactggccaactctacaagaagacacacagatcgcacatgctgcgaagggcagaatgg 1320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1151 AGCTTCTGTGTATCTTTCAAAATGACCATGTTATGCAAAAAGAGGCGCTTGTT 1210

QY 1321 cgcagatcgtgactatctgtccacaagaaggtgtctcgggtgactcgtatataccctcct 1380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1211 CTG-----TTTTCANCAACCTTGGTTCCAGCGGTTCTCTGATGTGACTA 1255

QY 1381 tgaagtgtgctgttacacagcggccagcaatgaagagtgatgtgtgtcagcagcg 1440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1256 TTTCCAAACAGAGTTACAGTTCCGGGTGAGGATTTGGTGAAGTTTTCATGACGACTG 1315

QY 1441 tgaaggtgtgtcgtatggaatgtgctgttc 1473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1316 TTAGCGGACGCTGTGACTTACAGATTTCTATCC 1348

```

```

RESULT 7
US-08-470-702-1
Sequence 1, Application US/08470702
Patent No. 5631149
GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO

```

APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,702
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-470-702-1

Query Match 21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 1.6e-97;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY 121 ctgagactgagcgaatcgaatcattatctcctcagagatcgaatggaagcgg 180
DB 11 CTGATTAATGAGATCAACAGTCTATTATCAAAATTGTCACGTGACAGATTGCTAAGAACCG 70
QY 181 atggtgcaagactgcaactgttaatactcgcgatacgaataactgtgtggaacatgyc 240
DB 71 ATGTGTATACAAAGTCTCTCTCTTACACACAGAGATGACTTACTGTGAGGTCTCTTTC 130
QY 241 agggcaccacgaagaattggaactataccagaagaatgagcattcacaagccatcgtatca 300
DB 131 AAGGCATCATTAAGAAATTTGGATTACATCAAAAGATATGGCTTTACTGCTATTGGAATTT 190
QY 301 ccccgctlaagccagctgcccagacacacgcacatactgagaatgctcattacagtcgtcat 360
DB 191 CTCACGTTGTTGAATAACATTCGCGATTAACACACACATATGTTTACTGCTATACAGGTTTCT 250
QY 361 ggcagcaggtatatactctctgaagaanaaactagcgcaactgagatgactgaaagcgc 420
DB 251 GGATGAAGAACAATATCAAAATTAATGAATAATTTGGTACTGCTGATGATTTGAAGCTTT 310
QY 421 tctctcgagcccttcatagagagggagatattatgtgtcagtggtgtgtaaccata 480
DB 311 TGGCACAAGAAATTCAGATCGTGAATATGTTGTTAAATGGGATATCGTTTACCAACCAATT 370
QY 481 tgggtatgatgagcgggtagctcagtcagttacagtggttlaaacgcgttaagctccc 540

DB 371 ACGCAGTGAATGGCAGTGGAGATAGTATGATGATTACCTAGAGTACACCCGTTCAAGACC 430
QY 541 aagactacttccaccgctctcgttcaatcaaaactatgaaatcagactcaggtgag 600
DB 431 AAAGTACTTCATTAATCTAGTCTTATTTCAAACATATGATGACCAAGCTCAGGTTCAAA 490
QY 601 atgtctgctaggaagataaacacgtctcctcctgctgctatcgtatccaccacaagatgag 660
DB 491 GTTCCTGGGAAGGTGAGCTCTTCAGTTGCATTACCAATTTGAAACGGAAGATAGCGACG 550
QY 661 tcaagaatgaatgtagcagactgggtggatcatctgtatccgaactcatatgaagcc 720
DB 551 TGGCCTTCAGTTTCAATTTCTGGGTTAAAGATTTTGTGGCAATTTACTCAATTTGATGTT 610
QY 721 tccgtatcgaacagaataaacaacgctcagaagagactctgcccgtgtaacaaagccg 780
DB 611 TAAGAAATTTGATGCTGCTAAACATGTGACCAAGCTTTTCCGGATTTGTATGTCAT 670
QY 781 caggcgtgactgtatcggcgaggtgctcgaaggtgataccggtactactgttccctaac 840
DB 671 CTGAGATTACTAGTAGGCGAAGTTTCCAAAGAGACCCAGCTTATACATCCCATACC 730
QY 841 aagaagtcabtgagcggcgtactgaactatccattactatccactcccaagccctca 900
DB 731 AAATTTACATTTCCAGGGGTAGTAATTTATTCATTTGACTACCAACCAACGATTTTTA 790
QY 901 agtcaaccctccgagcagatggaagaccttacaacatgatacaacgctcaaatccgact 960
DB 791 AAATCTAGTATTCAAAGTTCCAGTGAATGACCTCAATTAATGATTTCAAGCTTGTCTCAGTT 850
QY 961 gtccagactcaaacacccctcgggacactcgtcgaagacagacacccaggttcgctt 1020
DB 851 GTTCGATCCAACTTTGTTGACAAATTTGTAGAAATATCCATTAATGAAGATTCGCTT 910
QY 1021 cttaacccaagacatagccctcgcgaagaacgtcgcgaattcatcctcacaagcag 1080
DB 911 CAATGACCAAGCCCAAAAGTTGATTTTAATGCTATTATGCTATTGCTTTGGGTGATG 970
QY 1081 gaatcccatcactcagcgcgcgcaagaacagacactaagccgagcggaaagaccccgca 1140
DB 971 GATTTCTGTCAATTTACTATGACAAAGAACAGGCTTGAACGGAATAAGTCAACCA 1030
QY 1141 accggaagcaacactggtcctcgcgctcagccagcagcagcagcgtgtacaaatgattg 1200
DB 1031 ACAGAGAGGCTTGTGGTATTCGGCTCAACAAAGAGGTGACTATTATCAAGCTCATTTG 1090
QY 1201 cctcgcggaagcgaatccggaactatgcatatagcaagaatacagatctgtaacctaca 1260
DB 1091 CCAAAAGCTAATGCTCCCAAGAAAGCCGCCGCTTTATCAAGACTCAAGCTATGCCACTCGC 1150
QY 1261 agaactggccatcacaagaagacacacaagactgcatgctcgcaagggcagagatggt 1320
DB 1151 ACCTTTCTGTGATCTTTTCAAAATGACCAATGTTATTGCAACAAAAGAGGCGCGCTGTTT 1210
QY 1321 ggcagatcgtgactactctgtccacaaggggtgctcgggtgattcgtatataccctcct 1380
DB 1211 CTG-----TTTCAACAACCTTGGTTCCAGCGGCTCTCTGATGTGACTA 1255
QY 1381 tgaagtgggtgggtttacacagccggtccagcaattgaaggaagtcattgctgcacagcg 1440
DB 1256 TTTCACAACAGAGTTACAGTTCCGGGTGAGGATTTTGGTAAGATTTTGACATGACGACTCG 1315
QY 1441 tgaacgttggttcgagatgaaatgtgctgttc 1473
DB 1316 TTAGCGGCAAGCTCTGACTTACAAAGTTTCTATCC 1348

RESULT 8
US-08-470-702-2
Sequence 2, Application US/08470702
Patent No. 5631149
GENERAL INFORMATION:

APPLICANT: MATSUI, IKUO
 APPLICANT: ISHIKAWA, KAZUHIKO
 APPLICANT: MIYAIRI, SACHIO
 APPLICANT: HONDA, KOICHI
 TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
 TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
 TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 STREET: 8110 Gatehouse Road, Suite 500 East
 CITY: Falls Church
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22042
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/470,702
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/204,656
 FILING DATE: 02-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: WEINER, MARC S.
 REGISTRATION NUMBER: 32,181
 REFERENCE/DOCKET NUMBER: 234-252P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 TELEX: 248345
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1404 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (synthetic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO

Query Match	21.6%;	Score 373.8;	DB 1;	Length 1404;
Best Local Similarity	55.5%;	Pred. No. 1.6e-97;		
Matches 751; Conservative	0;	Mismatches 587;	Indels 15;	Gaps 1.
QY 121	ctgcgcgactcggatccgacatccattatcttccttcacggatcgatttgcgaagacgg	180		
Db 11	CTGATTAATATGGAGATCACAGCTCTATTATTACCAATTTGTCTACACAGATTTCGTAGAACGG	70		
QY 181	atcggctcgcagcactcgcgacttgtaataactcgcgatacgaataacttggctggaacatg	240		
Db 71	ATGGTATATCAAGTGGCTTCCTGTAACACAGAAAGATGAGATTACTGTTGGTCTTTTCC	130		
QY 241	agggcatactcgcacaaagtgcgactatccaggaaatggcttcacaaagcatctgatca	300		
Db 131	AAGGCTTCCTAAAGAAAGTTGGATTACATCAAAAGATATGGGCTTTATCTGCTATTGGATT	190		
QY 301	cccccgttacaagccacgcgcgccccagaccacgcgcatactgagatgctctaccatg	360		
Db 191	CTCCAGTTGTTGAAACATTTCCCGATTAACACACAGCATATGGTATTCCTTATCATGTGTTG	250		
QY 361	ggcagcagatatatactctctcgaagaaacacgagcacttgcgaagctgttgaagggcg	420		
Db 251	GGATGGAAGAACATATTCAAAATTTAAAGAAAATTTGGTACTGCTGATGATTTAAGTCTT	310		
QY 421	tctcttcggcccttcaggaaggagatgattcttctatgtctcagatgtggttgcacacata	480		

Db	311	TGGACACAGAAATTGCACGATCGTGAATGTTGTTAATGCTGGATATGTTACCAACATT	370
Qy	481	tggacata tga tga agc agg tag ctc agc tgc gat taca cgt g ttt aaac cgt ttc agt tcc	540
Db	371	ACGGCACTGATGGCAGTGGAGATAGATAGCATTACTTCAGAGTACACCCCGTTCAACGACC	430
Qy	541	aagactacttccaccgcgttctgtttcatctcaaaactatgaatcagatcagatcagtttga	600
Db	431	AAAGTACTTCCATTAACACTGCTCTTAATTCAAACTATGATGACCAAGCTCAGGTTCAA	490
Qy	601	atgtcgttcagagataaacacgtctcctctgtccgatctcgtatataccaccaaga tgg	660
Db	491	GTTCTCTGGAGGAGACTCTTCAGTTGCAATTACCGAGTTTGAGAACGGACATAGCCACG	550
Qy	661	tcaagatgaatggttaacgttgg tgg a tca t t g t a t c g a a c t a c c a t t g a g g c c	720
Db	551	TGGCTCAGTTTCAATTCTTGGGTTAAAGATTGTTGGCAATTACTCAATTTGATGGTT	610
Qy	721	tcgctatcgacacagtaaaacacgctccagaaggactctgcccgggtatcaacaagccg	780
Db	611	TAAGATTTGATAGTCACTAACTAGTGGACCAAGGCTTTTCCCGGATTTGTTAGTGCAT	670
Qy	781	cagcgtgtactgtatcggcgagg tgc tgc agc ggt g a t c c g g c c t a a c t t g c c a c c	840
Db	671	CTGGAGTTACTCAGTAGGGGAAGTTTCCAAAGAACCCACCTTATACATGCCATATCC	730
Qy	841	agaagctatcgacggcggtactgaactlcccaattactatcacctctcaagccctca	900
Db	731	AAATTTACATTCCAGGGGTTAGTATTTATTCATTGTATACCCACACGAGATTTTTA	790
Qy	901	agtcacacctccggcagatgagacgacctacaacatgatacaacgctcaaatccagat	960
Db	791	AAACTACTGATTCAAGTTCCAGTAGTAGTCACTCAATGATTTTCAAGCGTGTCTTCCAGTT	850
Qy	961	gtccagactcaaacactccctggcacatctgtcgaagaacacagacacccacagttcgtt	1020
Db	851	GTTCCGATCCAACTTTGTTGCAAACTTTGTAAATAATCAATATGAAAGGTTCCGTT	910
Qy	1021	cttaccacaagacatagccctcgccaagaagctgcgaacatcatalcatcctcaacgag	1080
Db	911	CAATACCAACGACCAAGATTGATTTCTATGCTATTCGATTTTGCTTTGGGTGATG	970
Qy	1081	gaatcccatcatcagccggccaagaacagactacgcccggcggaaacgaaccccgcga	1140
Db	971	GATTTCTGTCATTACTATGACACAAAGACGGCTGAGCGCAAAAMAGTGACCCAAACA	1030
Qy	1141	accgcgaagcaacctggtctccggtaccaccgcgcagacagggcgctgcagaatgaatt	1200
Db	1031	ACAGAGAGCGCTTGGTGGTTATCCGGCTTACAACAAGAGAGTGTACTTTCAAAGCTCATTTG	1090
Qy	1201	ccttcgcgaagcgaatccgcgaacatctgcaltgacaagaatacagaattcgtgactaca	1260
Db	1091	CCAAAGCTATATCTGCCGAAGACGCCGCCGCTTATATCAAGCTCAAGCTATGGCACCTGGC	1150
Qy	1261	agaactgcccatactacaagaacgaacaagatcgcgcatctgcgaaggcacagatgggt	1320
Db	1151	AGCTTCTGTGATCTTTTCAATGACCATGTTATTCGAACAAAAGAGGACAGCTGTGTTT	1210
Qy	1321	cgcagatcgtgactctctgtccacaagaaggtgcttcgggttggtgtataccctcct	1380
Db	1211	CTC-----TTTTCAACAACCTTGGTTCCACCGGCTTTCTTCGATGTGACTA	1255
Qy	1381	tgaagtgtgcgggttacaacgcgcgcgcgaacatgtagcggaggtcattgctgcgaacgc	1440
Db	1256	TTTCCACACAGGTTACAGTTCGGTGAGGATTGGTAGAAGTTTGATGACATGAGTACAG	1315
Qy	1441	tgaagttgttcgga tgg a a t g t g c c t g t c 1473	
Db	1316	TTAGCGCAGCTGTGACTTACAAAGTTCTATCC 1348	

Sequence 3, Application US/08470702
Patent No. 563149
GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE.
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,702
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-470-702-3

Query Match 21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 1.6e-97;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY 121 ctgagcgagctgagctgcaatccatttattctctccacgagatcttggcaagagcg 180
DB 11 CTGATTAATGAGATCAAGTCTTATTAATCAAAATGTCACGAGATTGTCAGAACCG 70
QY 181 atggctgacgactgactgttaatactcgagatcagaataactgtgtgtggaacatgyc 240
DB 71 ATGCTGATACAAAGTCTTCCTGATACACAGAAAGATACACTTACTGTGCTGCTTTTCC 130
QY 241 agggacatcatcgacaagtgtgactatactcagggaaatggtgcttcacagccatctgatca 300
DB 131 AAGGCATCATAAAGAACTTGGATTACATCAAAAGATATGGCTTTACTGCTATTGATTT 190
QY 301 ccccgctacagccagctgcccagagccacccgataatgagatgcttaccatgctact 360
DB 191 CTCGAGTTGTTGAATAATTCCTCGATACACAGCATATGTTATGCTTATCATGCTCTCT 250
QY 361 gcaagcagagatatatactctctgaacgaanaactcagcactgcagatgacttgaagcgc 420
DB 251 GGATGAAGACATATACAAAATTAATGAAAACTTGTGTACTGCTGATGATTTGAAGCTTT 310

QY 421 tctcttgcgccctcatgagagggagtgatcttattgctcagatgctgtgctaacata 480
DB 311 TGGCAACAAGAAATGCGACGATCGATGATGTTAAATGATGATGATTCACCAACCAT 370
QY 481 tggcctatgagagcgagtgagctcagtgattacagtgctgtttaaaccgttaagctcc 540
DB 371 ACGCACTGATGACGATGGAGATAGATGATTCGATTCAGAGTACCCCGCTTAACGACC 430
QY 541 aagactcttccaccgcttctgttcaatcaaaactatgaaatcagaactcagatgagag 600
DB 431 AAAGTACTTCCTCAATACTACTGCTTATTCAACACTATGATGACCAAGCTCAGTTCAAA 490
QY 601 atgtctgcaagagagataaacaactgtctctctcctatctcgalaccacaagatgag 660
DB 491 GTTGCTGGGAAGTGATCTCTTCAGTTGATTCACCAATTTGAAAGCAAGATACGACG 550
QY 661 lcaagaatgaaatgtaagactggtgagatcatctgtatcagactctcatatgagcgc 720
DB 551 TGGCTTCAGTTTCAATTCCTGGGTTAAAGATTTTGTGGCAATTACTCAATTTGATGTT 610
QY 721 tccglatcgacacaglaaaacacgctccagaagactctgcccgggtacaaacaaagccg 780
DB 611 TAAGAATTTGATAGTCTAAACATGTGACCAAGCTTTTCCGGATTTGTTAGTCAAT 670
QY 781 caggcgtgactgctatcgcgagagtgctcgaagtgatcgcgcctacactgttccctacc 840
DB 671 CTGGAATTTACTCAGTAGGCGAAGTTTCCAAAGAGACCCAGCTTATATCATCCCATACC 730
QY 841 agaagctcatgagacgagctgactaactatccattaccatccactccacagcccttca 900
DB 731 AAATTTACATTCACAGGGGTTAGTAATATTCATTCATTCATTCATTCATTCATTCAT 790
QY 901 agtcaacctcgcgcagcatgagacactctacaatgataatgatacaacacgctcaactc 960
DB 791 AAACCTACTGATTCAGATTCAGAGTGTGATTCATCAATGATTTCAAGCGTTGCTTCAG 850
QY 961 gtccagactcaaacctccctggcgacactcgtcgaagacacagacacacagaggttcgtt 1020
DB 851 GTTCGATTCACAACTTTGTTGACAACTTTGTGAAATATCAGATATGAAAGTTCCGCTT 910
QY 1021 cttaaccaacagacatagccctcgcgaagaacgtcgaacatctcatcctcacaagcgc 1080
DB 911 CAATGACACAGCCAAATTTGATTTCTAATGCTAATTCATTTGCTTTGGGTGATG 970
QY 1081 gaatcccatcatctacgcgcgcgaagaacagcactacgcgcgcgcgcgcgcgcgcgcgc 1140
DB 971 GATTCCTGCTATTACTATGACAAAGAACAGCTTGACGGAAGAAAGTGACCCAAACA 1030
QY 1141 accggaagcaacctgagctcctgagctccagcagcagcagcagcagcagcagcagcagc 1200
DB 1031 ACAGAGAGGCTTGTGCTTATTCGGCTACACAAAGAGGTGATATTAACAGCTCATTTG 1090
QY 1201 cctcgcgaacgcaatccggaactatgcatatgaaagatacagagatctgtactcata 1260
DB 1091 CCAAAAGCTAATGCTGCGAAGAACGCCGCTTATCAAGACTCAAGGATGACACCTGCG 1150
QY 1261 agaactcgcacctctacaaagacgacaaacagatcgcacatgcgaagggcagaatggt 1320
DB 1151 ACCTTTCTGTGATCTTTCAAAATGACATGTTAATGCAAAAGAGGACGCTGTTT 1210
QY 1321 cgcagatcgtgactatctgtccaaagaggtgctcggagatctgctatccctctcct 1380
DB 1211 CAG-----TTTCAACAAACCTTTGGTTCCAGCGGTTCTTCTATGATGACTA 1255
QY 1381 tgaagtgtgctgttacacagcgcgcgaacttgaacggaagatgctgtcgcagcgcgc 1440
DB 1255 TTTCCAAACACAGTTACAGTTCCGCTGAGATTTGCTGAGAGTTTGCATGACGTA 1315
QY 1441 tgaagctgtgtgagatggaatgctgtc 1473
DB 1316 TTAGCGGACGCTGACTTACAAAGTTTCTATCC 1348

RESULT 10
US-08-467-831-1
Sequence 1, Application US/08467831
Patent No. 5635378
GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: MATSUI, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,831
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-467-831-1

Query Match 21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 1.6e-97;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY 121 ctgcgacgtggcgatcgatcattattctctctcaacgagtcatttgaagagc 180
DB 11 ctgattaaatggagatcacagctatttattcaantttctctctacacagattttctgaacccg 70
QY 181 atggtgcgacgtcgactgttaatactgcgatacagaataactgttggaacatggc 240
DB 71 atggtgatcaaaatgctctctctgaacacagaaagatagactttactgtgtgtctttcc 130
QY 241 agggcatcatcgacaagtggatgataatcagaaggaatgggtctcaacgcatctgatac 300
DB 131 aaagcatcatcaaaagagttgatattacataaaagatattggcctttactgtatttggattt 190
QY 301 ccccggttaacggccagctgcccacagaccacgcacatagagatgctaccatgact 360
DB 191 ctccagttgttgaaacatttcccgatatacagacagatattgttattgcttattcattgctttct 250

QY 361 ggcacgagatatatactctcgaacgaaaactacgacatcgacatgacttgaagcgc 420
DB 251 ggaatgaagaacaatattcaaaatttatgaaaactttggctactgattgattgaagcttt 310
QY 421 tctcttcgaccttcataagaggggagatcattatgttgatgtgtgtgtacaccta 480
DB 311 tggcacagaaatggacagatcggatattgttttaagtggtgatattcttaccacactt 370
QY 481 tgggtatgatagagcggttagctcagtcattacagttgtttaaccgcttcagctcc 540
DB 371 acggcagatgatggcagatgagatagatattgatttaccacagatgacacccgttcaacgacc 430
QY 541 aagactactccaccgcttctgttcttcaataaactaagaatacagactcagttgag 600
DB 431 aaagatcttccatgactgactgcttatttcaaatgattgatgaccagctcaggttcaaa 490
QY 601 attgctgcttaagagataaacgtctctcttgcctgctatctgataccacaggaatgtg 660
DB 491 gttcctgggaaggtgactcttcagttgcatattaccagatttgagaaacggaatagcgacc 550
QY 661 tcaagaaatgaaatgtaacgactgggtgagatcattgtgtatcgaactactcattgaagcc 720
DB 551 tggcctcagttttcaattcttggggttaaaagattttgttgacatttctcaattgatgatt 610
QY 721 tccgtatcgaacaaagtaaaacacgctccagaagactctggccgggtatcaacaaagccg 780
DB 611 taagaaattgatacgttaaaacatgatggaccacagctttttcccgattttgttattgcatt 670
QY 781 cagcggtactctatcgtgaggtgctgacggtgatccggctcactgtctccatcc 840
DB 671 ctggagtttactcgtatggcgaaagttttccaaagagaccacgctttatcattgcccattacc 730
QY 841 aagaagtcacgagacgctcactgaactacatccattacttccactcctcaacgcttca 900
DB 731 aaatttcaattccagggggttagtaaatattatccattgactaccacacagagatttttta 790
QY 901 agtcaacctcggcagcatgaaagacacctcaacaactatgatacaacacgyltaactccgact 960
DB 791 aaacttactgattcaagttccagttcagttgactcaaaagatttcaaggtgcttccagtt 850
QY 961 gtccagactcaacactcctcctggacatctcgtcgaagaacacagaaacccaggttcgctt 1020
DB 851 gtttgagctccactttgttgacaaactttgttagaaatcagaaatgaaagagttgcctt 910
QY 1021 cttaacacaacagatagccctcgcgaagaacgctgcgacatcattcatcctcaacgacg 1080
DB 911 caatgacacagcagaccaaaagttgatttctaatgctattgcatattgttgggtgatg 970
QY 1081 gaatcccatcatcatcagcggcggaagaagaacacactacgcgcggcggaagaaaccccgga 1140
DB 971 gtaattcctgctatttactatngacaaagaaacaaagccttgagcggaamaaagtgacccaaca 1030
QY 1141 acccggaagcaacctggctcgggtctaacccgacacagcagcgtctaaagaattatg 1200
DB 1031 acagagagcccttgggttattccgcttacaacaaagagatgactatttacaagctcattg 1090
QY 1201 cctcgcgaacgcaatccggaactatgccaattagcaaaagatagacgattcgttgactaca 1260
DB 1091 ccaaaagctaaatggccagaaacgcccggcttttcaaaagctcaagctatggccacctgcg 1150
QY 1261 agaactggcccatctacaagaagacacacaaacagatcgccatgacgaaggaacagataggt 1320
DB 1151 agctttctgctgattttttccaatgacacattgttttgcaaaaaaagagacgcttggttt 1210
QY 1321 cgcagatcgtaactactctgccaacaaggggtgctcgggtgattcgtatataccctcct 1380
DB 1211 ctg-----ttttcaaaacaccttggttgacagcggttctcttcatgattgacta 1255
QY 1381 tgaatgtgcgggttacaacagcggcgacgaatgaaaggaggtcattgtgcgcagcgcg 1440
DB 1256 ttttccaaacacaggttaccggttccggtgaggaatttggtaagattttgacatgacgactcgt 1315
QY 1441 tgacggttggttcggaatgaaatgtgcctgttc 1473

```

Db 1316 TTAGCGGACAGCTCTGACTTACAGTTTCTATCC 1348
RESULT 11
US-08-467-831-2
Sequence 2, Application US/08467831
Patent No. 5635378
GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,831
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
HYDROTHERMAL: NO
ANTI-SENSE: NO
US-08-467-831-2

Query Match 21.6%; Score 373.8; DB 1; Length 1404;
Best Local Similarity 55.5%; Pred. No. 1.6e-97;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

QY 121 ctgcgactgagcagatcgatccatcttcttccttcacagcagcttgcaaggcgg 180
Db 11 CTGATTAATGAGATCAGATCTTATTTATCAATTTGTCACAGATTTGCTTGAACCG 70
QY 181 atgggtcgaagcagctgagctgttaatactcgcgatacagaatactgtgtggaacatgac 240
Db 71 ATGGTGAATACAAAGTCTCTGTAACACAGAAATAGACTTACTGTGGTGGTTCTTCC 130
QY 241 agggcatcatcgacaagtgtgacatataccaggaatgggcttccacgcatctgataca 300
Db 131 AAGGATCATTAAGAAAGTTGGATTACATCAAAAGATATGGCGCTTACTGCTATTTGGATT 190
QY 301 ccccgctacagccagctgcccagaccagccatctgagatgagctacacatgagctact 360

Db 191 CTCAGTTGTTGAANAACATTCGCCATACACAGATATGGTTATGCTTATCATGTTGGT 250
QY 361 ggcagcagatatactctctcgaagaaactacagcagctagagatgacttgaagcgc 420
Db 251 GGATGAAGAACATATACAAATTAATGAAAACTTGGTACTGCTGATGATTTGAAGCTT 310
QY 421 tctctcgccctcctacagagaggggagtgatctctatagtcgagtggtgtgtaaccata 480
Db 311 TGGCAACAAGAAATTCACAGATCGTATGTGTTAAAGGGATATCGTATCCACCAATT 370
QY 481 tggcctatgagagcgagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 540
Db 371 ACGCAGATGATGCGAGTGGAGATAGTATGATTAATGATGATGATGATGATGATGATGAT 430
QY 541 aagactatccaccgcttctgttcaatcaaaactaagaatgagatcagactaagttgaag 600
Db 431 AAAAGTACTTCCATTAACATCTGCTTATTTCAAACTGATGATGATGATGATGATGATGAT 490
QY 601 atgctgctagagagataacacatgctctcctctgctcagtcagtcagtcagtcagtcagtc 660
Db 491 GTTCTGGAAGGATGATCTTTCAGTTGCTATTCACAGATTTGACAGAGAGAGATGAGCAGC 550
QY 661 tcaagaatgaatgagcagcagctggtggtggtggtggtggtggtggtggtggtggtggt 720
Db 551 TGGCTCAGTTTCAATTTCTGGGTTAAAGATTTGTTGGCAATTTACTCAATTTGATGTT 610
QY 721 tcgctacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 780
Db 611 TAAAGATTTGATAGTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670
QY 781 cagcgtgctatgctatgctgctgctgctgctgctgctgctgctgctgctgctgctgct 840
Db 671 CTGAGTTTACTCAGTATGAGGAGGATTTTCCAGAGAACCCAGCTTATATGCGCCATATC 730
QY 841 agaacgtcatggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 900
Db 731 AAAATTCATTTCCAGGGGTTAGTAATTAATTCATTTGATGATGATGATGATGATGATGAT 790
QY 901 agtcaacctccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
Db 791 AAATCTGATTTCAAGTTCAGTCAAGTTCAGTCAAGTTCAGTCAAGTTCAGTCAAGTTCAGT 850
QY 961 gtccagactcaaacctcctgagcagcagcagcagcagcagcagcagcagcagcagcagc 1020
Db 851 GTTCGATCCAACTTTGTGACAACTTTGTGACAACTTTGTGACAACTTTGTGACAACTTTGT 910
QY 1021 cttacaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
Db 911 CAATGACCAAGCAGCAAAAGTTGATTTCTAATGCTAATTCATTTGCTTTGGGTGATG 970
QY 1081 gaatcccatcatctacgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140
Db 971 GTATTCCTGTGATTTCTATGAGCAAGCAAGCTTTGAGCGGAAAAATGACCAAAACA 1030
QY 1141 accggaagaaacctggtcctcgggctacccgagcagcagcagcagcagcagcagcagc 1200
Db 1031 ACAGAGAGGCGCTTGTGTTATCCGCTTCAACAAGAGAGTACATTAACAGTCAATTCG 1090
QY 1201 cctcggagaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1260
Db 1091 CCAAGCTTAATGCTGCAAAAGCGCGCTTATCAAAAGCTAAGCTAATGCAACCTCGC 1150
QY 1261 agaacctgcccactcaacaagcagcagcagcagcagcagcagcagcagcagcagcagc 1320
Db 1151 AGCTTCTGTGATCTTTCAATGACATGATTTATGCAACAAAGAGGACGCTTGTGTT 1210
QY 1321 cgcagactgctgactatctgtccacaagagtgctcgggtgagtcgtatataccctcct 1380
Db 1211 CTG-----TTTCAACAACCTTGTTCCAGCGGTTCTTGTGATGACTA 1255
QY 1381 tgaagtgtggttatacaagccgagcagcagcagcagcagcagcagcagcagcagcagc 1440
```



```
OY 1381 tgaagtgtcgagggttacacagccgcccgaattgacgaggtcatgtgctgcacgaccg 1440
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1256 TTTCACACACAGGTTACAGTTCGGGTGAGGATTGGTACAGATTGGTACAGTCACTG 1315
OY 1441 tgaaggtgtgttcgagatggaatgtgctgttc 1473
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1316 TTAGCGCAGCTCTGACTTACAAGTTTCTATCC 1348

RESULT 13
US-09-189-060B-63
; Sequence 63, Application US/09189060B
; Patent No. 6270968
; GENERAL INFORMATION:
; APPLICANT: Dalboge, Henrik
; APPLICANT: Sandal, Thomas
; APPLICANT: Kauppinen, Markus
; APPLICANT: Borge, Diderichsen
; TITLE OF INVENTION: Method Of Providing No. 6270968e1 DNA Sequences
; FILE REFERENCE: 4772.204-US
; CURRENT APPLICATION NUMBER: US/09/189,060B
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/DK97/00216
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-09-189-060B-63

Query Match 4.0%; Score 68.6; DB 4; Length 93;
Best Local Similarity 94.7%; Pred. No. 1.4e-10;
Matches 71; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 41 gaagcattatgatgcgcgtgtgtgtctctattctgtacggccttcagtcgcgcgcg 100
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 gatccattatgatgcgcgtgtgtgtctctattctgtacggccttcagtcgcgcgcg 63
OY 101 acctgcttgctgc 115
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 acctgcttgctgc 78

RESULT 14
US-09-189-060B-67
; Sequence 67, Application US/09189060B
; Patent No. 6270968
; GENERAL INFORMATION:
; APPLICANT: Dalboge, Henrik
; APPLICANT: Sandal, Thomas
; APPLICANT: Kauppinen, Markus
; APPLICANT: Borge, Diderichsen
; TITLE OF INVENTION: Method Of Providing No. 6270968e1 DNA Sequences
; FILE REFERENCE: 4772.204-US
; CURRENT APPLICATION NUMBER: US/09/189,060B
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/DK97/00216
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 919
; TYPE: DNA
; ORGANISM: Hybrid
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(919)
```

```
US-09-189-060B-67

Query Match 3.9%; Score 68.4; DB 4; Length 919;
Best Local Similarity 98.6%; Pred. No. 5.5e-10;
Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 46 cattatgatgtcgcgtgtgtgtctctattctgtacggccttcagtcgcgcgcacctg 105
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 cattatgatgtcgcgtgtgtgtctctattctgtacggccttcagtcgcgcgcacctg 62
OY 106 ctltgctgc 115
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 ctltgctgc 72

RESULT 15
US-09-189-060B-65
; Sequence 65, Application US/09189060B
; Patent No. 6270968
; GENERAL INFORMATION:
; APPLICANT: Dalboge, Henrik
; APPLICANT: Sandal, Thomas
; APPLICANT: Kauppinen, Markus
; APPLICANT: Borge, Diderichsen
; TITLE OF INVENTION: Method Of Providing No. 6270968e1 DNA Sequences
; FILE REFERENCE: 4772.204-US
; CURRENT APPLICATION NUMBER: US/09/189,060B
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/DK97/00216
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 922
; TYPE: DNA
; ORGANISM: Hybrid
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(922)
US-09-189-060B-65

Query Match 3.9%; Score 68.4; DB 4; Length 922;
Best Local Similarity 98.6%; Pred. No. 5.5e-10;
Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 46 cattatgatgtcgcgtgtgtgtctctattctgtacggccttcagtcgcgcgcacctg 105
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 cattatgatgtcgcgtgtgtgtctctattctgtacggccttcagtcgcgcgcacctg 61
OY 106 ctltgctgc 115
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 ctltgctgc 71
```

Search completed: June 13, 2002, 18:24:33
Job time: 13242 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 13:40:45 ; Search time 1884.09 Seconds
(without alignments)
12421.763 Million cell updates/sec

Title: US-09-710-339-1

Perfect score: 1734
Sequence: 1 tcacatcaagctctccttc.....aaaaaaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estha:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_iny:*
15: em_gss_pln:*
16: em_gss_vtc:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	256.8	14.8	583	9	AA786071	AA786071 j4h07a1.r
2	240.2	13.9	541	9	AA787396	AA787396 n3a12a1.r
3	227.2	13.1	561	9	AA783132	AA783132 c1c11a1.r
4	210.2	12.1	464	9	AA965792	AA965792 o5e03a1.r
5	189.8	10.9	466	9	AI211322	AI211322 o6c07a1.r
6	150.8	8.7	347	9	AA785831	AA785831 h8a04a1.r
7	143.2	8.3	376	9	AA786814	AA786814 m6c11a1.r
8	140.2	8.1	374	9	AA785147	AA785147 g5c10a1.r
9	139.4	8.0	382	9	AA785574	AA785574 g8h01a1.r
10	138.6	8.0	358	9	AI209683	AI209683 c1c11a1.r
11	133	7.7	389	9	AI212282	AI212282 x1f04a1.r
12	131.8	7.6	358	9	AA784386	AA784386 d5d11a1.r
13	131.2	7.6	304	9	AI210530	AI210530 i7h08a1.r
14	116.2	6.7	405	9	AA787924	AA787924 r2h03a1.r
15	112.2	6.5	362	9	AA787426	AA787426 n3c10a1.r
16	111.6	6.4	293	9	AA785338	AA785338 g8d04a1.r
17	104	6.0	283	9	AA788570	AA788570 a1f06c9.r

18	103.8	6.0	548	10	BG278743	BG278743 a6h08np.r
19	96.6	5.6	335	9	AI210822	AI210822 10e09a1.r
20	83.2	4.8	450	9	AI210074	AI210074 g9b03a1.r
21	75.6	4.4	876	10	BM400915	BM400915 5009-0-80
22	74.6	4.3	287	9	AA785414	AA785414 g7b08a1.r
23	74.6	4.3	402	12	A2877703	A2877703 #06 3.0 M
24	72.6	4.2	668	9	AU060851	AU060851 AU060851
25	72.6	4.2	677	9	AU060198	AU060198 AU060198
26	71.4	4.1	232	9	AI209785	AI209785 c9f06a1.r
27	69.4	3.5	559	9	AU052395	AU052395 AU052395
28	60.4	3.4	293	9	AA901823	AA901823 NCM981173
29	59.4	3.4	353	9	AI210266	AI210266 h4b03a1.r
30	58.8	3.4	653	9	AU039903	AU039903 AU039903
31	58.8	3.4	695	9	AU034096	AU034096 AU034096
32	57	3.3	457	9	AA787302	AA787302 n1e09a1.r
33	55.2	3.2	500	12	B67199	B67199 CpG0015B Cp
34	54.8	3.2	681	12	CNS02E0D	AI133990 "retrodon
35	54.4	3.1	525	9	AI544417	AI544417 SD10962.5
36	53.6	3.1	608	10	BI887904	BI887904 ZF637-1-0
37	53.2	3.1	645	9	AI389106	AI389106 GH20192.5
38	53	3.1	738	12	PT008M15U	AL447629 Parameciu
39	52.8	3.0	649	10	BI641907	BI641907 SD25571.5
40	52.2	3.0	703	12	AQ162044	AQ162044 mgxb0010F
41	52.2	3.0	895	12	CNS0071A	AL066286 Drosophila
42	51.2	3.0	511	10	BM004465	BM004465 TgESTz98
43	51.2	3.0	922	12	CNS0073W	AL066784 Drosophila
44	51.2	3.0	587	10	BE446309	BE446309 WHE1455_H
45	51	2.9	626	10	BF473952	BF473952 WHE0839_C

ALIGNMENTS

RESULT 1
LOCUS AA786071 583 bp mRNA linear EST 31-JUN-1998
DEFINITION j4h07a1.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone j4h07a1 5', mRNA sequence:
AA786071.1 GI:2846239

ACCESSION AA786071.1 GI:2846239
VERSION EST.
KEYWORDS Emericella nidulans.
SOURCE Emericella nidulans.
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiata; Trichocomaceae; Emericella.

REFERENCE 1 (bases 1 to 583)
Kupfer,D., Gray,U., Hausner,J., Lai,H., Martin,W., Aramayo,R., Prade,R. and Roe,B.
An Aspergillus nidulans EST Database

JOURNAL Unpublished (1998)
COMMENT other_ESTS: j4h07a1.fl
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu

We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center
Seq primer: SK
High quality sequence stop: 526.
Location/Qualifiers
1..583
/organism="Emericella nidulans"
/strain="FGSC A26"
/db_xref="taxon:162425"
/clone="j4h07a1"
/clone_lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library"
/tissue_type="vegetative mycelia, asexual structures"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:

FEATURES
source

Query Match 14.8%; Score 256.8; DB 9; Length 583;
 Best Local Similarity 65.2%; Pred. No. 1.9e-29;
 Matches 378; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762
 Email: broe@ou.edu
 We anticipate the future release of the cDNA clones to the Fungal
 Genetics Stock Center
 High quality sequence stop: 510.
 Location/Qualifiers
 1..541
 /organism="Emicella nidulans"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="n3a12a1"
 /clone_11b="Aspergillus nidulans 24hr asexual
 developmental and vegetative cDNA lambda zap library"
 /tissue_type="vegetative mycelia, asexual structures"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
 3' end of cDNA cloned into XhoI site of pBluescript"

620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762
 Email: broe@ou.edu
 We anticipate the future release of the cDNA clones to the Fungal
 Genetics Stock Center
 High quality sequence stop: 510.
 Location/Qualifiers
 1..541
 /organism="Emicella nidulans"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="n3a12a1"
 /clone_11b="Aspergillus nidulans 24hr asexual
 developmental and vegetative cDNA lambda zap library"
 /tissue_type="vegetative mycelia, asexual structures"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
 3' end of cDNA cloned into XhoI site of pBluescript"

Query Match 13.9%; Score 240.2; DB 9; Length 541;
 Best Local Similarity 65.2%; Pred. No. 6.2e-27;
 Matches 353; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

620 Parrington Oval, Norman, OK 73019, USA
 Tel: 405 325 4912
 Fax: 405 325 7762
 Email: broe@ou.edu
 We anticipate the future release of the cDNA clones to the Fungal
 Genetics Stock Center
 High quality sequence stop: 510.
 Location/Qualifiers
 1..541
 /organism="Emicella nidulans"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="n3a12a1"
 /clone_11b="Aspergillus nidulans 24hr asexual
 developmental and vegetative cDNA lambda zap library"
 /tissue_type="vegetative mycelia, asexual structures"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
 3' end of cDNA cloned into XhoI site of pBluescript"

RESULT 2
 LOCUS AA783132 541 bp mRNA linear EST 31-JUL-1998
 DEFINITION n3a12a1.r1 Aspergillus nidulans 24hr asexual developmental and
 vegetative cDNA lambda zap library Emicella nidulans cDNA clone
 n3a12a1 mRNA sequence.
 ACCESSION AA783132
 VERSION AA783136.1 GI:2847627
 KEYWORDS EST.
 SOURCE Emicella nidulans.
 ORGANISM Emicella nidulans.
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; Emicella.
 1 (bases 1 to 541)
 Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R.,
 Prade, R., and Roe, B.
 An Aspergillus nidulans EST Database
 Unpublished (1998)
 Other_ESTs: n3a12a1.f1
 Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
 Department of Chemistry and Biochemistry
 Advanced Center for Genome Technology, University of Oklahoma

RESULT 3
 LOCUS AA783132 561 bp mRNA linear EST 29-JUL-1998
 DEFINITION c1c1a1.r2 Aspergillus nidulans 24hr asexual developmental and
 vegetative cDNA lambda zap library Emicella nidulans cDNA clone

accession AA783132.1 GI:2843300
 version AA783132.1
 keywords EST.
 source Emericella nidulans.
 organism Emericella nidulans.
 reference 1 (bases 1 to 561)
 authors Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R.,
 title An Aspergillus nidulans EST Database
 journal Unpublished (1998)
 comment Other ESTs: c1c1a1.f1
 contact: Bruce A. Roe, University of Oklahoma, broeou.edu
 department of Chemistry and Biochemistry
 advanced center for genome technology, University of Oklahoma
 620 Parrington Oval, Norman, OK 73019, USA
 tel: 405 325 4912
 fax: 405 325 7762
 email: broeou.edu
 we anticipate the future release of the cDNA clones to the Fungal
 Genetics Stock Center
 seq primer: SK
 high quality sequence stop: 511.
 location/qualifiers
 1. 561
 /organism="Emericella nidulans"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="c1c1a1"
 /clone_1ib="Aspergillus nidulans 24hr asexual
 developmental and vegetative cDNA lambda zap library"
 /tissue_type="vegetative mycelia, asexual structures"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
 3' end of cDNA cloned into XhoI site of pBluescript"
 BASE COUNT 124 a 163 c 147 g 127 t
 ORIGIN
 Query Match 13.1%; Score 227.2; DB 9; Length 561;
 Best Local Similarity 68.1%; Pred. No. 5.5e-25;
 Matches 316; Conservative 0; Mismatches 148; Indels 0; Gaps 0;
 Db 129 tggcgatcgacatcattatcttctctccgcatgatttgcaagagcgatgctg 188
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 98 TGGCGACGCGATTCCTCTCTCGACGAGCGATTGCTCGCAGCGAGGATCG 157
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 189 acgactgacattgtaactcgatcaagaactcgtgtgtaacatgacagcgc 248
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 158 ACGACCGTGGCTGTGATCTGGCTCAACGAGATACGCGCGCAGCTGGCAGGCGATC 217
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 249 atcgacaagtgtgactatataccaggaatgggttcacagcacttgatcaccccg 308
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 218 ATCAACGAGCTTGAT 277
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 309 acagcccgatgcccccgacacgcgcataatgagatgctcctcattgctatgcaacg 368
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 278 ACCGCGAGATTCCCGATGTACCGCTGTGGAACGGGCTTCATGGTACTGGCAAG 337
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 369 gataatactctctgaagaagaactcgaatgcaatgactggaagcgctctctcg 428
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 338 AACATATACGGTGTGACACCAACCTGGGACACCGACGATATACAGGCGCTGTGCGAG 397
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 429 gcccttcataagagggatgatatgctatgctgctgctgctgctgctgctgctgctgct 488
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 398 GCGGTCAATGATCGGGGATGATCTGCTGATGCTGCTGATGCTGCTGCTGCTGCTGCTG 457
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 489 gatgagagggatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 548
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 458 GCGGCGCGCGGATGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 517

Oy 549 ttccaccgctgttcttcaataaactatgaatcagactca 592
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 518 TTCCATTCGTACTGCGCATATCAACATGATGACATCAGTGCA 561
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 RESULT 4
 AA965792 464 bp mRNA linear EST 31-Jul-1998
 LOCUS 05603a1.r1
 DEFINITION Aspergillus nidulans 24hr asexual developmental and
 vegetative cDNA lambda zap library Emericella nidulans cDNA clone
 05603a1.5', mRNA sequence.
 accession AA965792
 version AA965792
 keywords EST.
 source Emericella nidulans.
 organism Emericella nidulans.
 reference 1 (bases 1 to 464)
 authors Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R.,
 title An Aspergillus nidulans EST Database
 journal Unpublished (1998)
 comment Contact: Bruce A. Roe, University of Oklahoma, broeou.edu
 department of Chemistry and Biochemistry
 advanced center for genome technology, University of Oklahoma
 620 Parrington Oval, Norman, OK 73019, USA
 tel: 405 325 4912
 fax: 405 325 7762
 email: broeou.edu
 we anticipate the future release of the cDNA clones to the Fungal
 Genetics Stock Center
 seq primer: 73
 high quality sequence stop: 453.
 location/qualifiers
 1. 464
 /organism="Emericella nidulans"
 /strain="FGSC A26"
 /db_xref="taxon:162425"
 /clone="05603a1"
 /clone_1ib="Aspergillus nidulans 24hr asexual
 developmental and vegetative cDNA lambda zap library"
 /tissue_type="vegetative mycelia, asexual structures"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
 3' end of cDNA cloned into XhoI site of pBluescript"
 BASE COUNT 118 a 135 c 94 g 117 t
 ORIGIN
 Query Match 12.1%; Score 210.2; DB 9; Length 464;
 Best Local Similarity 65.9%; Pred. No. 2.1e-22;
 Matches 305; Conservative 0; Mismatches 158; Indels 0; Gaps 0;
 Db 820 cggcctacactgttcctaccagaacgltcagagcgctgactgaactatccattact 879
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1 CGCGATATACCTGCCCTTACACAGATTATATGACGGGCTCATGAACCTACCCATATATT 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 880 atcaactctcaagccttcaagtaactcggcgatgagatgagacccctcaacaatga 939
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 61 ACCCCCTTCGATGGCTTCACTCTCGACGCGGAGATGCGATCTCTATACATGA 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 940 tcaaacgctcaaatccgactgtccagactcaacactctcgggacatccgctcgagaacc 999
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 121 TCAACACAGTGGCTCAAAATGTGCGGATCTTACACTGCTTGGAACTTATTCGAGAAC 180
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 1000 acgacaaccacggttcgctcttcaacaaagacatgacctcgccaagacgttcgacg 1059
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 181 ATGACATCTCTGATTCGCCAATATATCTCGGATATGATGATGATGATGATGATGATG 240
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Oy 1060 catcaatcaatcccaagagaatcccatcatcctcgaacgagcgaagaagcagctag 1119
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 241 CGTTCTCTTCTTGACGAGCAATCCCTATGTTATGCGGCGCAGAGCAACACTATT 300

QY	1120	ccggcggaagcagcaccgccggaacgcggaagcaactcgtgctctcgtggctacccgacgaca	1179
Db	301	CAGCGAGCATGATGCCCTTATTAACCCGGGACCCGGTTGGTGTCCTCTACTCGACAGCT	360
QY	1180	gcgaagctgacaaatgaattatgctcctcgcggaagcaaacctcgcgaactatgcatagcaaa	1239
Db	361	CAGAGCTATACAAAGTTCACTCGGACCACTTAAACAAGATCGGAATACTGGGCATTCCAAAG	420
QY	1240	atacaggaatcgtagacctaacgaagaatgycatcatcaacaaga	1282
Db	421	ATTCCAGTTATCTCACTCTCCGGAATATCTCTTTTACAGCGA	463
RESULT	5		
LOCUS	AI211322	466 bp	mRNA
DEFINITION	06c07a1.r1 Aspergillus nidulans 24hr asexual developmental and		
	vegetative cDNA lambda zap library Emericella nidulans cDNA clone		
	06c07a1 5', mRNA sequence.		
ACCESSION	AI211322		
KEYWORDS	AI211322.1 GI:3773264		
SOURCE	EST.		
ORGANISM	Emericella nidulans.		
	Emericella nidulans		
	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;		
	Eurotiales; Trichocomaceae; Emericella.		
	1 (bases 1 to 466)		
REFERENCE	Kuper,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R.,		
AUTHORS	Prade,R. and Roe,B.		
TITLE	An Aspergillus nidulans EST Database		
JOURNAL	Unpublished (1998)		
COMMENT	Other ESTs: 06c07a1, fl		
	Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu		
	Department of Chemistry and Biochemistry		
	Advanced Center for Genome Technology, University of Oklahoma		
	620 Parrington Oval, Norman, OK 73019, USA		
	Tel: 405 325 4912		
	Fax: 405 325 7762		
	Email: broe@ou.edu		
	We anticipate the future release of the cDNA clones to the Fungal		
	Genetics Stock Center		
	Seq primer: T3		
	High quality sequence stop: 321.		
FEATURES	Location/Qualifiers		
source	1..466		
	/organism="Emericella nidulans"		
	/strain="FGSC A26"		
	/db_xref="taxon:162425"		
	/clone="06c07a1"		
	/clone_1ib="Aspergillus nidulans 24hr asexual		
	developmental and vegetative cDNA lambda zap library"		
	/issue_type="vegetative mycelia, asexual structures"		
	/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:		
	XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript		
	3' end of cDNA cloned into XhoI site of pBluescript"		
BASE COUNT	102 a 129 c 129 g 106 t		
ORIGIN			
Query Match	10.9%; Score 189.8; DB 9; Length 466;		
Best Local Similarity	70.4%; Pred. No. 2.4e-19;		
Matches 254; Conservative	0; Mismatches 107; Indels	0; Gaps	0;
QY	129	tgagcagcagcatcattatctctctcgtcagcagtcgatttgcaagcagcagatggcg	188
Db	104	TGGCGCACCCAGTCATTAATTTCTCTGACGGACGATTCGTCGACGACGAGATCG	163
QY	189	acagactcgacttgaatcacycgagatcagaataactgtgtggaacatgcagcagc	248
Db	164	ACGACCGCTGGCTGTGATCTGGCTCAACGAGATATCTCGCGGACGCTGGCAGGACATC	223
QY	249	atcgacaagttggactatataccaggaatggcgctcacagcacttgatcacaccgct	308

Db	224	ATCACACAGCTGATTATATATCCAAAGCATATGGGATGTGCATCGCAATCTGGATTACACTTACC	283
QY	309	aagcgccagctgccccagacaccgatattgtagatgccttacatgtcctaacyggacgacg	368
Db	284	ACCGACAGATGTTCCCATGTGTCACCGGTGTGGGAACGGGCTTACATGGCTACTGGCAGAAG	343
QY	369	galataactctctgaagaanaactacgycgatcagatcagatcttgaaggcgctcttcg	428
Db	344	AACATATACGGTGTCAGACACAACCAACTGGGACACGCCGACAGATATCATGGGCTGTGCGAG	403
QY	429	gcccttcacagagaggggatgtatctcttaagtgcgatgtgttgccaacatatggcat	488
Db	404	GCGCTCCATGATCGGGGCAATGTAATTCATGCTGGAGTGTTCGCAACCAATGTTCTTAT	463
QY	489	g	489
Db	464	G	464
RESULT	6		
LOCUS	AA785831	347 bp mRNA linear EST J1-JUL-1998	
DEFINITION	hba04al.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone hba04al.5', mRNA sequence.		
ACCESSION	AA785831		
VERSION	AA785831.1	GI:2845999	
KEYWORDS	EST.		
SOURCE	Emericella nidulans.		
ORGANISM	Emericella nidulans. Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; Emericella. 1 (bases 1 to 347) Kupfer,D., Gray,J., Hausner,T., Lal,H., Martin,W., Aramayo,R., Prade,R. and Roe,B. An Aspergillus nidulans EST Database Unpublished (1998)		
TITLE	JOURNAL		
COMMENT	Other-ESTs: hba04al.f1 Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu Department of Chemistry and Biochemistry Advanced Center for Genome Technology, University of Oklahoma 620 Parrington Oval, Norman, OK 73019, USA Tel: 405 325 4912 Fax: 405 325 7762 Email: broe@ou.edu We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center Seq primer: SK		
FEATURES	Source	High quality sequence stop: 323. Location/Qualifiers 1..347	
		/organism="Emericella nidulans" /strain="FGSC A26" /db_xref="taxon:162425" /clone="hba04al.1"	
		/clone_lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library" /issue_type="vegetative mycelia, asexual structures" /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; 5' end of cDNA cloned into EcoRI site of pBluescriptb 3' end of cDNA cloned into XhoI site of pBluescript"	
BASE COUNT	84 a	91 c	72 g 87 t 13 others
ORIGIN			
Query Match	Best Local Similarity	8.7%; Score 150.8; DB 9; Length 347;	
Matches	Conservative	0; Mismatches 125; Indels 0; Gaps 0;	
QY	775	aagcgcaagcgctgactgtatcgcgcgagtgctgcacggtgatacgcgacctaatgtc	834
Db	4	ATNCCNTNGGCTAACNCNMTNGATAAATCTTCATGGGAGACCGCGCATACTGCC	63

QY	835	ccaccagaagcgtacatgaagcgcgtactacgaacttccatttacttacttccacttcctcaag	894
Db	64	CTTNNNAAATTTATATGAGCAGGGGTCAGAACTACCCCATATATTACCCCTTCTGATG	123
QY	895	ccttcaagtcacactcgcgcagcatalgagagacctatcacatgatacaaacgylcaaat	954
Db	124	CGTTCAAGTCTCCGAGCGGGAGCATGTGGGATCTCTATACATNANGAGAGATCGGCT	183
QY	955	cgcagatgcacagactccacacatctcgtggcacaattcgtatgagacacagacaacacagyt	1014
Db	184	CAAAATGTGCGGATCCTACACTGCTTGGAAACTTTATGAAAAACCATATACAAATCTTCAT	243
QY	1015	tgccttcttacaacaaagacacatagccctcgcgacaagaaagyltcgcagcatlcatlctca	1074
Db	244	TCCCCAACTATACTCGGGATATGAAATCGGGCCAAAAGAGTCTCGTTCCTTCTGTGA	303
QY	1075	acgacggaatccccatcatlctacgcgcgcgcaagaacgacacia	1117
Db	304	CCGACGGAAATCCCTATTTGTTATATGAGGCGCAGAGAACCAACTCA	346

RESULT	7
AA786814	
LOCUS	
DEFINITION	
AA786814	376 bp mRNA linear EST 31-JUL-1998
m6c11a1.r1	Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans CDNA clone m6c11a1 5', mRNA sequence.

ACCESSION	AA786814	GI:2847144
VERSION	AA786814.1	
KEYWORDS	EST.	
SOURCE	Emricella nidulans.	
ORGANISM	Emricella nidulans	

TITLE Eukariota, Fungi, Ascomycota, Pezizomycotina, Eurotiomycetes,
REFERENCE Eurotiiales; Trichocomaceae; Emericella.
AUTHORS 1 (bases 1 to 376)
 Kuper, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R.,
 Prade, R. and Roe, B.
TITLE An *Aspergillus nidulans* EST Database
JOURNAL Unpublished (1998)
COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu

Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: SK
High quality sequence stop: 367.

FEATURES	source	Location/Qualifiers
	1..376	
		/organism="Emmericella nidulans"
		/strain="PGSC A26"
		/db_xref="taxon:162425"
		/clone="mbc1a1"
		/clone_11b="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library"
		/tissue_type="vegetative mycelia, asexual structures"
		/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript 3' end of cDNA cloned into XhoI site of pBluescript"
BASE COUNT	92 a	98 c 91 g 95 t
ORIGIN		

Query Match	8.38;	Score 143.2;	DB 9;	Length 376;
Best Local Similarity	65.1%;	Pred. No. 2.5e-12;		
Matches 242;	Conservative	0;	Mismatches 128;	Indels 2;
				Gaps 2

615 gataacactctctcccttgctgatctcgatccaccacaagatgtgtcaagaatgaatgg 674

Db	2	GACCATTAAGTCTCCTTGACTGATCTGAACACCCAGACTCCGAGGTACGCACTCTGG	61
Qy	675	taagactgggtggatcattgtatcgaactaccatccatgcgcccgcgatcgcagaca	734
Db	62	TATGATCGGTATCAAGGACATCGTTCCCAATTACTGTGGGATGGGTCGCCATTGGATTCC	121
Qy	735	gttaaaacag-tccagaagaactctctggccgggtacacacaagccg-cagcgcttact	792
Db	122	GTCAAGCACCTTTGACAAATGATTTCTGGCCGGGTTATTTATATCCCGACTGGGCTACA	181
Qy	793	gtatccgcgaggtgcctcgacggtgacccgcctaacactgtccctacagaaacgtatg	852
Db	182	GCATCGSTGAAATCTACCATGSGGGAACCGGAGTTATACCTGGCCTTTACGAGATTATGTGG	241
Qy	853	acgcgctactgaaactatcccatctactatccaactccctcaacgccttaagtcaactccg	912
Db	242	ACGGGGTCATGAACTAACCCATAATATTAACCCCTTCTAATGCGCTTCAAGTCCCTCGAGCG	301
Qy	913	gcagcatagcgcgcctctacaacatgatacaaacgcttaaatcgcagctgtccgaactcaa	972
Db	302	GGACCAATGTCGGATCTCTAACAATGATCAACACAGTGCCTCAAAATTGTGGGATCTTA	361
Qy	973	caactccctggaca	984
Db	362	CACGTCTTGGA	373

LOCUS	DEFINITION
AA785147	374 bp mRNA linear EST 28-JUL-1996
LOCUS	g5c10a1.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library
DEFINITION	g5c10a1.5', mRNA sequence.

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
AA785147	AA785147.1	GI:2845315	EST.	<i>Emericella nidulans</i> .	
				<i>Emericella nidulans</i> .	
				Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes	
				Eurotiales; Trichocomaceae; Emericella.	
				(phases 1 to 374)	

AUTHORS	Knapke, D., Gray, J., Hausner, J., Lal, H., Matlack, W., Almaguer, R., and Roe, B.
TITLE	An <i>Aspergillus nidulans</i> EST Database
JOURNAL	Unpublished (1998)
COMMENT	Other_ESTs: g5c10a1..f1

Contact: Bruce A. Roe, University of Oklahoma, broeou@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broeou@ou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: SK
High quality sequence stop: 340.

FEATURES	source
Location/Qualifiers	1. .374
/organism="Emmericella nidulans"	
/strain="FGSC A26"	
/db_xref="taxon:162425"	
/clone="g5c10a1"	
/clone_11b="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library"	
/tissue_type="vegetative mycelia, asexual structures"	
/note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2: XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript 3' end of cDNA cloned into XhoI site of pBluescript"	
BASE COUNT	84 a 107 c 99 g 83 t 1 others
ORIGIN	

Query Match	Similarity	8.1%	Score 140.2	DB 9	Length 374
Best Local	Similarity	71.3%	Pred. 71.2e-12		
Matches	184	Conservative	0	Mismatches 74	Indels 0
					Gaps 0
QY	129	tggcgatcgcgaatccattatttcctctctcgcagatcgaattgtgcaaggacgagatgctc	188		
Db	114	tggccacnancacgctccatctatcttccctcgacggagccgattgctgcgcacgagcagatcg	173		
QY	189	acgaactgcgaacttgaaactctgagatcagaataactctgtgtggaacatggagagcattc	248		
Db	174	acgacccgtgcgctgtgactctgcgtcgaacggagatattctgcggcgccagcctggcagggcattc	233		
QY	249	atcgacaagttygactatataaccagggaatgggcttcacaagccatctygataccccgctt	308		
Db	234	atcAACCCAGctTGAATTATATCCAAACATAGGgATTACtGCATCTGGATTACACCTTATC	293		
QY	309	acaagccacgtcgcccccagaccacgcgcatatygagatgctctacacggctcattcggcagacg	368		
Db	294	ACCGAGCAGATTCGCCGATGTCAACGCCCTTGTGGAACGGCCTTCATAGGCTACtGGCAGAAg	353		
QY	369	gatatataactctctgaaac	386		
Db	354	AACATATACGGGTGTGAC	371		

RESULT	9	
LOCUS	AA785574	
DEFINITION	AA785574 382 bp mRNA linear EST 31-Jul-1998	
ACCESSION	g8b01a1.1	Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone
VERSION	g8b01a1.5	' mRNA sequence.
KEYWORDS	AA785574	
SOURCE	AA785574.1	GI:2845742
ORGANISM	EST.	
	Emericella nidulans.	
	Emericella nidulans	
	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;	
	Eurotiales; Trichocomaceae; Emericella.	
REFERENCE	1 (bases 1 to 382)	
AUTHORS	Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R., Prade, R. and Roe, B.	
TITLE	An Aspergillus nidulans EST Database	
JOURNAL	Unpublished (1998)	
COMMENT	Contact: Bruce A. Roe, University of Oklahoma, broeeu.edu Department of Chemistry and Biochemistry Advanced Center for Genome Technology, University of Oklahoma 620 Parrington Oval, Norman, OK 73019, USA Tel: 405 325 4912 Fax: 405 325 7762 Email: broeeu.edu We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center Seq primer: SK High quality sequence stop: 359. Location/Qualifiers 1..382	
FEATURES		
source		

```

/organism="Emmericella nidulans"
/strain="RGSC A26"
/db_xref="taxon:162425"
/clone="g8h01a1"
/clone_1fb="Aspergillus nidulans 24hr asexual
developmental and vegetative cDNA lambda zap library"
/tissue_type="vegetative mycelia, asexual structures"
/notes="Vector: pBluescript SK-; Site1: EcoRI; Site2:
XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
3' end of cDNA cloned into XhoI site of pBluescript"
BASE COUNT
87 a 110 c 102 g 83 t

```

Query Match	8.0%;	Score 139.4;	DB 9;	Length 382;
Best Local Similarity	69.9%;	Pred. No. 9.4e-12;		
Matches 188;	Conservative	0;	Mismatches 81;	Indels 0;
				Gaps 0;

OY	129	tcgcacgcagcatccattattcttcctcttcacggaatcgatttcgaagacagatgctgc	188
Db	114	TGCGCCAGCCAGCTGGATCTATTCTCTCTGCAGCACCGGATTCGCTGCACAGGACGATCG	173
OY	189	acgactgcgacttglylaaactcgcgatacgaanaaactcgttgcgtgaatatgcacgggcac	248
Db	174	ACGACCGCTGCTGTGATCTGCTGCCTCAACGGGAGATCTCGCGGCGGACGCTGGCAGGCGATC	233
OY	249	atgcacaagctggagcctatccagagaaatggcttcacagcactcttgatcacaccccgctt	308
Db	234	ATTCACCAAGCTTGATTTATATTCACAGACATGGGATTCATCTCCATCTGGATTTAACCTATC	293
OY	309	acagaccacagctgcgccccagaccacgcgcatalgtagagatgctcctacatgctactgcagcag	368
Db	294	ACCAGAGCGATTTCCCGATATGTCACCGCTGTGGAAACGGGCTTCATGAGCTACTGGCAGAA	353
OY	369	gatatatactctcgaacggaanaactaagg	397
Db	354	AACATATACGGTGTGCAOCCACCACTTGGG	382

RESULT 10	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
AI209683	358 bp mRNA linear EST 19-OCT-1998	cllclal.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone cllclal 5', mRNA sequence.	AI209683			AI209683.1 GI:3771625	EST. Emericella nidulans. Emericella nidulans.	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; Emericella.	Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R., Prade, R., and Roe, B.	An Aspergillus nidulans EST Database	Unpublished (1998)	Other ESMs: cllclal.fl
										Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu		
										Department of Chemistry and Biochemistry		
										Advanced Center for Genome Technology, University of Oklahoma		
										620 Parrington Oval, Norman, OK 73019, USA		
										Tel: 405 325 4912		
										Fax: 405 325 7762		
										Email: broe@ou.edu		
										We anticipate the future release of the cDNA clones to the Fungal		
										Genetics Stock Center		
										Seq primer: 73		
										High quality sequence stop: 305.		
										location/Qualifiers		
										1..358		

```

/organism="Emericella nidulans"
/strain="FGSC A26"
/db_xref="taxon:162425"
/clone="c1c1a1"
/clone_1ib="Aspergillus nidulans 24hr asexual
developmental and vegetative cDNA lambda zap library"
/tissue_type="vegetative mycelia, asexual structures"
/note="vector: pbluescript sr-; Site_1: EcoR; Site_2:
XhoI; 5' end of cDNA cloned into EcoRI site of pbluescript
3' end of cDNA cloned into XhoI site of pbluescript"
BASE COUNT
79 a 102 c 98 g 79 t

```

Query Match	8.0%;	Score 138.6;	DB 9;	Length 358;
Best Local Similarity	72.3%;	Pred. No. 1.3e-11;		
Matches 180;	Conservative 0;	Mismatches 69;	Indels 0;	Gaps 0;

Oy 129 tggcgatcgcaatccattattctcttcacgcgatcgatttcgaagcagcgtggtcg 188

Db	108	TTGGGGCAACCGATGATCTATTTCCTCTGACGGACGATTCGCTCCGACGGAGGATCG	167
Qy	189	acgactgcgacttgtaactgcgatacaagaataactgtgtgtgaacatgycagggatc	248
Db	168	ACGACCGCTGCTTGATCTGGCTCAACGGAGAAATCTCGCGGAGCTGGCAGGGCATC	227
Qy	249	atcgacaagtgtgagcatatccagggaatgagcttccaaagcaactgatatcaacccglt	308
Db	228	ATCAACCAAGCTTGATTTATTCAGACACATGGAGATTCACTGCCATCTGATATACACTATC	287
Qy	309	acagcccgagctgcgccccagcaaccgcgcatatgtgagatgctctcaatggtactgtgcaag	368
Db	288	ACCGAGCGAGATTCCTCGATGTACACCGCTTGTGGAAACGGCTTCATGGCTACTGGCAGAG	347
Qy	369	gatatatac	377
Db	348	AACATATAC	356

RESULT	11
AI12282	
LOCUS	389 bp mRNA linear EST 19-OCT-1998
DEFINITION	x10441.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emeritella nidulans cDNA clone x10441.5', mRNA sequence.

ACCESSION	AI12282	GI:3774024
VERSION	AI12282.1	
KEYWORDS	EST.	
SOURCE	Emericella nidulans.	
ORGANISM	Emericella nidulans	

REFERENCE	1 (bases 1 to 389)
AUTHORS	Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R., Prade, R. and Roe, B.
TITLE	An Aspergillus nidulans EST Database
JOURNAL	Unpublished (1998)
COMMENT	Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu

Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel.: 405 325 4912
Fax: 405 325 7762
Email: brc@ou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: 73
High quality sequence stop: 307.

```
FEATURES      Location/Qualifiers
source        1. .389
```

```

/organism="Emerizella nidulans"
/strain="FGSC A26"
/db_xref="taxon:162425"
/clone="x1f04a1"

```

ORIGIN	BASE COUNT
clone_1lb="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library" /tissue_type="vegetative mycelial, asexual structures" /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript 3' end of cDNA cloned into XhoI site of pBluescript"	95 a 116 g 82 g 96 t

Query Match	7.7%;	Score 133;	DB 9;	Length 389;
Best Local Similarity	58.9%;	Pred. No. 8.5e-11;		
Matches 229; Conservative	0;	Mismatches 160;	Indels 0;	Gaps 0

Dy 1062 ttcatcattccccaacgaagaatccccatatctaacgcgagcccaagaacagcatacgc 1121
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 TTCTCCTTTTGTGACCGAGGATCCATTGGTTATCCGGGCCAGGAGCACACTATTCA 60

QY	1122	gqcgaaacgcagcccggaacccgcgaacgcgaagcaacccgctcggctcgggtaccgaccgacgcg	1181
Db	61	GGCACCATTGATGCTCTTATPACCGGGAGCCGGTTTGSTGTCTCTTACTCGACACAGCTCA	120
QY	1182	gagcgtcaacagttlaattgctccgcgcgaacggaatccgcgaactatgcatatgaagaat	1241
Db	121	GAGCTATTCAAATGTTTCATCGCGCACCTTAACAAGATCCCAAAAGCTGGCATTTCCAAAGT	180
QY	1242	acagattccgtaactaagaagacgcgcacatcctacaagaagacagacaaagatgcgcatg	1301
Db	181	TCCAGTTTATCTACTTCCCGGAAATCTCTTTTACAGCGATPAGCAACTATTTGCGCATG	240
QY	1302	cgcgaagggcacagatlyggtcgcagaatcgtgactatctgtccacaagglygtctcgagt	1361
Db	241	CCCAAGGGCTCTGGGGGGCTCAGGTCTCCATCTCCATCCACAAATATCGGCACCAATATC	300
QY	1362	gatttgtaaacctctcccttgagtgatgtaggggttacaagcgcgcgcagaaatctgacgag	1421
Db	301	GGTTCTTATACATTGCACTCTTATGATCATCTGGATATCAACAGGGGGGCCAACCTAATGGAA	360
QY	1422	gtcatatgctgcagacgcgttgaaagtttg	1450
Db	361	CTGTACACATGCTCTCTCTGTCCAGGTCGG	389

RESULT	12		
AA784386			
LOCUS	AA784386	358 bp	mRNA
DEFINITION	d5B11a1.r1 Aspergillus nidulans 24hr asexual developmental and		linear EST 29-JUL-1996

ACCESSION	vegetative cDNA lambda zap	library	<i>Emeticella nidulans</i>	cDNA clone
VERSION	d5b1a1.5'', mRNA sequence.			
KEYWORDS	AA784386			
SOURCE	AA784386.1	GI:2844554		
ORGANISM	<i>Emeticella nidulans</i> .			
	<i>Emeticella nidulans</i>			

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukariyota; Fungi; Ascomycota; Pezizomycotina; Eurotiiales; Eurotiomycetes; Eurotiaceae; Trichocommataceae; Emmentella.
1 (bases 1 to 358)
Kupfer, D., Gray, J., Hausner, J., Lai, H., Martin, W., Aramayo, R., Prade, R. and Roe, B.
An Aspergillus nidulans EST Database
Unpublished (1998)
Contact: Bruce A. Roe, University of Oklahoma, broeou.edu

Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broeou.edu
We anticipate the future release of the cDNA clones to the Fungal
Genetics Stock Center
Seq primer: SK
High quality sequence stop: 348.

FEATURES	Location/Qualifiers
source	1. .358

```

/organism="Emicicella nidulans"
/strain="RGSC A26"
/db_xref="taxon:162425"
/clone="d5blat"
/clone_1b="Aspergillus nidulans 24hr asexual
developmental and vegetative cDNA lambda zap library"
/tissue_type="vegetative mycelia, asexual structures"
/notes="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
3' end of cDNA cloned into XhoI site of pBluescript"

```

Query Match	7.6%	Score 131.8	DB 9	Length 358
Best Local Similarity	72.0%	Pred. No. 1.3e-10		
Matches 172	Conservative 0	Mismatches 67	Indels 0	Gaps 0

OY	129	tggcgatcgaacccatttatttccttcacaagatcagtttgcaaggacgagtgtgcg	188
Db	120	TGGCGAGCCAGCAGTCATCTATTTCCTCCGACGACCAGTTCCTGTGCACAGGAGGATTTG	179
OY	189	acgactgcgacttgttaataactcgcgatacagaatactgtgttgaaacaatgcgagcgatc	248
Db	180	ACGACCGCTGCCCTGATATGTGGCTCAACGAGATATCTGGCGGCGCAGCTGGCAGGCGATC	239
OY	249	atgcacaattgtgcatctaccaggaatgaggttacagccatctcgatacccccgctt	308
Db	240	ATCACACACTTGATTAATATTCACAAGACATGGGATTACGTCCCATCTGGATTTACACCTTAC	299
OY	309	acagcccaagctgccccaagacacccgatataagatgtagtacctaatgctcacgcagca	367
Db	300	ACCGACGATATTCOCGATGTCAACCGCTGTGGAAACGGGCTTCATGCTACTAGGAGAA	358

RESULT 13	
LOCUS	A1210530
DEFINITION	A1210530 304 bp mRNA linear EST 19-OCT-1998
VERSION	17h08al.r1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone
KEYWORDS	17h08al 5', mRNA sequence.
ACCESSION	A1210530
VERSION	A1210530.1 GI:3772472
SOURCE	EST.
ORGANISM	Emericella nidulans. Emericella nidulans. Eurotiota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; Emericella.
REFERENCE	1 (bases 1 to 304) Kupfer,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R., Prade,R. and Roe,B. An Aspergillus nidulans EST database Unpublished (1998)
AUTHORS	Other_ESTS: 17h08al.fl Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu Department of Chemistry and Biochemistry Advanced Center for Genome Technology, University of Oklahoma 620 Parrington Oval, Norman, OK 73019, USA Tel: 405 325 4912 Fax: 405 325 7762
JOURNAL	Email: broe@ou.edu
COMMENT	We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center Seq primer: T3 High quality sequence stop: 293. Location/Qualifiers 1..304 /organism="Emericella nidulans" /strain="FGSC A26" /db_xref="taxon:162425" /clone="17h08al" /clone_lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library" /tissue_type="vegetative mycelia, asexual structures" /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript 3' end of cDNA cloned into XhoI site of pBluescript"

FEATURES	
SOURCE	
BASE COUNT	
ORIGIN	

Query Match	7.6%	Score 131.2;	DB 9;	Length 304;
Best Local Similarity	64.5%;	Pred. No. 1,7e-10;		
Matches 196; Conservative	0;	Mismatches 108;	Indels	0; Gaps 0;

OY	945	accgtcaaatcgcgactgcacgaatacaactcctgtggacacatctcgtcagaaccagac	1004
Db	1	ACAGTCGCGCTCAAAATTGTCGGGATCTCACTGCTTGGAAACTTTATCGAGAACCATGAC	60
OY	1005	aaccaacggttcgctcttataccacaagacatagccctcgcgaagaacgtgcagatctc	1064

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 13:44:19 ; Search time 293.75 Seconds
(without alignments)
10134.909 Million cell updates/sec

Title: US-09-710-339-1
1734
1 tccatcaagctctccctc.....aaaaaaaaaaaaaaaa 1734

Sequence: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Scoring table:

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1734	100.0	1734	22	AAF90208
2	1707	98.4	1914	21	AAF12832
3	1063.8	61.3	1337	11	AAQ06786
4	806	46.5	1389	21	AAF11252
5	570	32.9	1956	16	AAQ88712
6	439	25.3	1389	21	AAF11252
7	378.6	21.8	1404	13	AAQ25079
8	377.4	21.8	4214	8	AAQ70916
9	375.4	21.6	1404	15	AAQ77668

10	375.4	21.6	1404	15	AAQ77665
11	374.2	21.6	483	21	AAF11314
12	373.8	21.6	1404	15	AAQ77667
13	373.8	21.6	1404	15	AAQ77666
14	369.2	21.3	2289	11	AAQ06388
15	368.8	21.3	605	21	AAF11273
16	349.6	20.2	2290	9	AAQ81525
17	348	20.1	2290	9	AAQ81477
18	344.2	19.9	1048	21	AAF11253
19	336.4	13.6	619	21	AAF12538
20	220.6	12.7	292	21	AAF11355
21	206.6	11.9	1602	21	AAF07664
22	190	11.0	1067	10	AAQ91274
23	190	11.0	1316	8	AAQ70507
24	190	11.0	1316	13	AAQ24569
25	190	11.0	1316	22	AAQ59332
26	190	11.0	1596	10	AAQ91273
27	178	10.3	1315	10	AAQ91074
28	165.4	9.5	592	21	AAF11268
29	143	8.2	3311	17	AAQ10562
30	143	8.2	3311	19	AAQ61458
31	143	8.2	7432	19	AAQ61459
32	129.2	7.5	173	21	AAF11472
33	92	5.3	1596	10	AAQ91273
34	70.6	4.1	839	21	AAF11277
35	68.6	4.0	93	19	AAQ15070
36	68.4	3.9	922	19	AAQ15072
37	68.4	3.9	922	19	AAQ15073
38	68.4	3.9	925	19	AAQ15076
39	68.4	3.9	928	19	AAQ15074
40	68.4	3.9	1438	10	AAQ91275
41	65.2	3.8	256	17	AAQ10566
42	63.4	3.7	915	19	AAQ15075
43	62.4	3.6	7417	23	AAQ16368
44	59.4	3.4	1067	10	AAQ91274
45	58	3.3	2582	18	AAQ73117

ALIGNMENTS

RESULT 1	
ID AAF90208	standard; DNA; 1734 BP.
XX	
AC AAF90208;	
XX	
DT 06-AUG-2001	(first entry)
XX	
DE Nucleotide sequence of a fungamyl-like alpha-amylose.	
XX	
KW Fungamyl-like alpha-amylose; glucanase; dextrinase; maltose;	
KW alcohol; starch; dough improver; brewing; starch liquification; ss.	
XX	
OS Aspergillus oryzae.	
XX	
FH Key	Location/Qualifiers
FT CDS	54..1550
FT	/*tag= a
FT	/product= "fungamyl-like alpha-amylose"
FT	/*tag= b
FT	114..1547
FT	/*tag= c
FT	
XX	
PN WO200134784-A1.	
XX	
PD 17-MAY-2001.	
XX	
PF 10-NOV-2000; 2000WO-DK00626.	
XX	
PR 10-NOV-1999; 99DK-0001617.	
XX	

Variant alpha amyL
Aspergillus niger
Variant alpha amyL
Variant alpha amyL
AMVI gene encoding
Aspergillus niger
Sequence encoding
5' end of alpha-am
Aspergillus niger
Aspergillus oryzae
Aspergillus niger
Fusarium venenatum
Plasmid NA2 DNA en
Sequence of the TA
Sequence of TAKA-
Aspergillus oryzae
Plasmid NA1 DNA en
Sequence of Asperg
Aspergillus niger
Alpha-amylose gene
Vector pBR70 conta
Aspergillus niger
Plasmid NA1 DNA en
Aspergillus niger
Humicola insolens
Hybrid DNA compri
Hybrid DNA compri
Hybrid DNA compri
DNA encoding Asper
Alpha-amylose gene
Hybrid DNA compri
Diosphila melanog
Plasmid NA2 DNA en
Actinoplanes sp. a

PA (NOVO) NOVOZYMES AS.
XX
PI Bisgard-Frantzen H, Svendsen A, Pedersen S;
XX WPI: 2001-367478/38.
DR P-PSDB: AAB84206.
XX
XX New variant of Fungamyl-1-like alpha-amylase, useful for production of
PT maltose syrups, includes mutations that improve stability against heat
PT and acidic pH.
XX
XX Example 1; Page 40-42; 49pp; English.
PS
XX
XX The present sequence encodes a fungamyl-1-like alpha-amylase. The
CC specification describes variants of this fungamyl-1-like alpha-amylase,
CC which have an alteration in one the amino acid regions 98-110, 150-160,
CC 161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion
CC or substitution of an amino acid or an insertion of an amino acid
CC downstream of a particular position. The variants retain alpha-amylase
CC activity, and have better heat stability and/or stability at acidic pH,
CC relative to wild-type enzyme. The variants can therefore be used at
CC higher temperatures (more efficient conversion or faster reaction, and
CC have reduced need for cooling and reduced risk of contamination). The
CC variants may also be used in conjunction with other enzymes,
CC particularly glucoamylase during dextrinisation. The variants are
CC used to produce syrups, particularly of high maltose content, or alcohol,
CC from starch; as dough improver for baked goods; in brewing, to increase
CC fermentability of the wort; and for liquefaction of starch.
XX
XX Sequence 1734 BP; 450 A; 460 C; 425 G; 399 T; 0 other:
SQ

Query Match 100.0%; Score 1734; DB 22; Length 1734;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tcaatcaagaactctccctctctgtgaacaataaaccacagaaggcattatgagtgcg 60
DB 1 tcaatcaagaactctccctctctgtgaacaataaaccacagaaggcattatgagtgcg 60
QY 61 cgtggtggtctattcttgtagcgccttcaggtgcgcgcctctgttggtgaagcg 120
DB 61 cgtggtggtctattcttgtagcgccttcaggtgcgcgcctctgttggtgaagcg 120
QY 121 ctcgcagactgcgcatcgcaatccatttattctctctcaagcagatcgattgcaagcagc 180
DB 121 ctcgcagactgcgcatcgcaatccatttattctctctcaagcagatcgattgcaagcagc 180
QY 181 atgggttcgaagactgcgactgttaatactgcggtcagaataactgtggtgaacatgyc 240
DB 181 atgggttcgaagactgcgactgttaatactgcggtcagaataactgtggtgaacatgyc 240
QY 241 aagggtacatcgacaagtgtgactatatacagggaatgggttcacagccatctgataca 300
DB 241 aagggtacatcgacaagtgtgactatatacagggaatgggttcacagccatctgataca 300
QY 301 ccccccgttacaagccagctgcgcacagaccagcatatggagatgctcaccatggtcact 360
DB 301 ccccccgttacaagccagctgcgcacagaccagcatatggagatgctcaccatggtcact 360
QY 361 ggcagcagagatatatactctctgaacgaaaactacgcgactgcagatgacttgaagcgc 420
DB 361 ggcagcagagatatatactctctgaacgaaaactacgcgactgcagatgacttgaagcgc 420
QY 421 tctcttcgcgccttcacgtagaggggagtatcttaatggtcgatggtgtgtaacata 480
DB 421 tctcttcgcgccttcacgtagaggggagtatcttaatggtcgatggtgtgtaacata 480
QY 481 tgggcatatagtagcggttagctcagtcgattacagtggttttaacccgttcagttccc 540
DB 481 tgggcatatagtagcggttagctcagtcgattacagtggttttaacccgttcagttccc 540
QY 541 aagactactcaccacccgtctcttcatccaactatgaagatcagctcaggttgagc 600

DB 541 aagactactcaccacccgtctcttcatccaactatgaagatcagctcaggttgagc 600
QY 601 attgctgctcagaagagataaacaactgctctctctgctcattctcgtataccaccaagagtgcg 660
DB 601 attgctgctcagaagagataaacaactgctctctctgctcattctcgtataccaccaagagtgcg 660
QY 661 tcaagaatgaatgtagcagcttggtggtgattcattggtlatcgaactactccatcgtagcgc 720
DB 661 tcaagaatgaatgtagcagcttggtggtgattcattggtlatcgaactactccatcgtagcgc 720
QY 721 tccgtatcgacaagaataaacaacgltccagaaggacttgcgcccgggtacaacaagccg 780
DB 721 tccgtatcgacaagaataaacaacgltccagaaggacttgcgcccgggtacaacaagccg 780
QY 781 caggcgtgactctatctgcgagagtgctctcgaagtgatccgcgctcacttgcctctacc 840
DB 781 caggcgtgactctatctgcgagagtgctctcgaagtgatccgcgctcacttgcctctacc 840
QY 841 agaacgctatggagcgggtactgaactatccattactactatccactccctcaagccttca 900
DB 841 agaacgctatggagcgggtactgaactatccattactactatccactccctcaagccttca 900
QY 901 agtcaacctccgcgcagcatgtagcagaccttacaacatgatcaaacacgcgtloaatccgact 960
DB 901 agtcaacctccgcgcagcatgtagcagaccttacaacatgatcaaacacgcgtloaatccgact 960
QY 961 gtccaagctaaactctctgtgagacattcgttgtagaacacagcaacccacggttcgct 1020
DB 961 gtccaagctaaactctctgtgagacattcgttgtagaacacagcaacccacggttcgct 1020
QY 1021 cttaacacaaacgacatagccctcgcgaagaacttcgaagacttcacatccatccctcaagcagc 1080
DB 1021 cttaacacaaacgacatagccctcgcgaagaacttcgaagacttcacatccatccctcaagcagc 1080
QY 1081 gaatcccatcatctacgcgcgcgaagaacagcactacgcgcgcgaagacccgcga 1140
DB 1081 gaatcccatcatctacgcgcgcgaagaacagcactacgcgcgcgaagacccgcga 1140
QY 1141 accgcgaagaacactgtgctctgcgtacccgcgcgaacagcagctgttacaagtattgc 1200
DB 1141 accgcgaagaacactgtgctctgcgtacccgcgcgaacagcagctgttacaagtattgc 1200
QY 1201 ccttcgcgaacgcacatccggaactatgcca ttagcaagaataaaggattcgttgaacctaca 1260
DB 1201 ccttcgcgaacgcacatccggaactatgcca ttagcaagaataaaggattcgttgaacctaca 1260
QY 1261 aagaactgcccacatcacaaagacgacaaacgactgcgcagtgcgaagggcacagatgggt 1320
DB 1261 aagaactgcccacatcacaaagacgacaaacgactgcgcagtgcgaagggcacagatgggt 1320
QY 1321 cgcagatcgtgactatcttctcaacaagggtgtcttcgggttgattcgtataccctctct 1380
DB 1321 cgcagatcgtgactatcttctcaacaagggtgtcttcgggttgattcgtataccctctct 1380
QY 1381 tgaatgtgtcgggtttacacagccgcgcgaagaattgacgcgaggttcatttgcgtcacagcg 1440
DB 1381 tgaatgtgtcgggtttacacagccgcgcgaagaattgacgcgaggttcatttgcgtcacagcg 1440
QY 1441 tgaacggttgcgtcggatggaatgtgctgttccatgycaggttggctacctaagggtat 1500
DB 1441 tgaacggttgcgtcggatggaatgtgctgttccatgycaggttggctacctaagggtat 1500
QY 1501 tgtatccgactcggagaattgtgcaggttagcagaagatcgttaagctcgttgaagggtgaga 1560
DB 1501 tgtatccgactcggagaattgtgcaggttagcagaagatcgttgaagggtgaga 1560
QY 1561 gtatataatgtagtactatccaactcgtgcattggaaagtgagatttgagattgata 1620
DB 1561 gtatataatgtagtactatccaactcgtgcattggaaagtgagatttgagattgata 1620
QY 1621 gttggaatcgttaactgtctgcatcccttataactctcgaattgttttcgaaacctaatg 1680
DB 1621 gttggaatcgttaactgtctgcatcccttataactctcgaattgttttcgaaacctaatg 1680

Db 1621 gttggagctgctactgctgcatcccttactcttcgattgttttcgaaccctaag 1680
 Oy 1681 ccaagcgcgtactgtctctatagaaaaaataaaaaaataaaaaa 1734
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1681 ccaagcgcgtactgtctctatagaaaaaataaaaaaataaaaaa 1734

RESULT 2
 AAF12832
 ID AAF12832 standard; cDNA; 1914 BP.
 XX
 AC AAF12832;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Aspergillus oryzae EST SEQ ID NO:5355.

XX Multiple gene expression; filamentous fungal cell; EST;
 KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
 KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
 KW culture condition; environmental stress; spore morphogenesis;
 KW metabolic pathway engineering; catabolic pathway engineering; ss.
 OS Aspergillus oryzae.
 XX
 PN WO200056762-A2.
 XX
 PD 28-SEP-2000.

XX 22-MAR-2000; 2000WO-US07781.
 XX
 PR 22-MAR-1999; 99US-0273623.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Berka RM, Rey MW, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;
 XX
 DR WPI; 2000-594572/56.
 XX

PT Monitoring differential expression of genes in filamentous fungal cells
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a
 PT substrate of expressed sequence tags -
 PT
 PS
 PS Claim 88; Page 2231-2232; 3161pp; English.

CC The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring
 CC the global expression of genes from FF cells allows the production
 CC potential of the microorganisms to be improved. New genes may be
 CC discovered, possible functions of unknown open reading frames can be
 CC identified and gene copy number variation and stability can be
 CC monitored. The expression of genes can be used to study how FF cells
 CC adapt to changes in culture conditions, environmental stress, spore
 CC morphogenesis, recombination, metabolic or catabolic pathway
 CC engineering. Using ESTs provides several advantages over genomic or
 CC random cDNA clones including elimination of redundancy as one spot on an
 CC array equals one gene or open reading frame, and organization of the
 CC microarrays based on function of the gene products to facilitate
 CC analysis of the results. AAF07478 to AAF1247 represents ESTs from
 CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
 CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
 CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
 CC all specifically claimed in the present invention.
 XX
 SQ Sequence 1914 BP; 484 A; 510 C; 461 G; 459 T; 0 other;

Query Match 98.4%; Score 1707; DB 21; Length 1914;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 tcaacataagcctccctctcttgaaacataaaccaccagaagcattatgagtgcg 60
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 3 tcaacataagcctccctctcttgaaacataaaccaccagaagcattatgagtgcg 62
 Oy 61 cgtgtgtgctctattcttgtaagcctcaagtcgagccctcttgctgctgaagc 120
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 63 cgtgtgtgctctattcttgtaagcctcaagtcgagccctcttgctgctgaagc 122
 Oy 121 ctgcgagctgagcagatccgaatccattatctctctcaagcagatgattgcaagcag 180
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 123 ctgcgagctgagcagatccgaatccattatctctctcaagcagatgattgcaagcag 182
 Oy 181 atggttcgaacgagctgagctgtaatactgagcagcaaaatagctgtgtgaagatgc 240
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 183 atggttcgaacgagctgagctgtaatactgagcagcaaaatagctgtgtgaagatgc 242
 Oy 241 agggacatacgaagaattggaactatccagaagatggcttcaacgaactgtgatca 300
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 243 agggacatacgaagaattggaactatccagaagatggcttcaacgaactgtgatca 302
 Oy 301 ccccgcttacaagccagctgcccagaccacagcatalgagatgcttcaatggtctact 360
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 303 ccccgcttacaagccagctgcccagaccacagcatalgagatgcttcaatggtctact 362
 Oy 361 ggcagcagatataatactcttcgaacgaactagcgacctgcaatgacttggaagcgc 420
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 363 ggcagcagatataatactcttcgaacgaactagcgacctgcaatgacttggaagcgc 422
 Oy 421 tctcttcgccccttcgaagagggatgatacttaatgctgagtgtgtgtaaccata 480
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 423 tctcttcgccccttcgaagagggatgatacttaatgctgagtgtgtgtaaccata 482
 Oy 481 tgggcatatgagcgggtagctcagtcgattacagtggtttaaacggttcagttccc 540
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 483 tgggcatatgagcgggtagctcagtcgattacagtggtttaaacggttcagttccc 542
 Oy 541 aagactacttccaccgcttctgtttatccaacatgaataagactcagtggttgg 600
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 543 aagactacttccaccgcttctgtttatccaacatgaataagactcagtggttgg 602
 Oy 601 attgctgagtagagaataacacgctctccttcgctgattcgcataccacaagatgag 660
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 603 attgctgagtagagaataacacgctctccttcgctgattcgcataccacaagatgag 662
 Oy 661 tcaagaatgattgtagcagctggtggtggtatcattggtatcgaaactcattcgacgc 720
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 663 tcaagaatgattgtagcagctggtggtggtatcattggtatcgaaactcattcgacgc 722
 Oy 721 tcggttcgacagaagtaaaacacgcttcagaagagacttcggccgggtgtaacaagcgc 780
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 723 tcggttcgacagaagtaaaacacgcttcagaagagacttcggccgggtgtaacaagcgc 782
 Oy 781 caggcgttactgtatcgcgcagagtgctcgacggtgatccggccttactgttcctacc 840
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 783 caggcgttactgtatcgcgcagagtgctcgacggtgatccggccttactgttcctacc 842
 Oy 841 agaagctatgagcggcgttactgtaactatccattactatccactcctcaacgcctta 900
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 843 agaagctatgagcggcgttactgtaactatccattactatccactcctcaacgcctta 902
 Oy 901 agtcaacttcggcagacatgagcagctcttaacaagatgtaaacacgctcaactcagact 960
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 903 agtcaacttcggcagacatgagcagctcttaacaagatgtaaacacgctcaactcagact 962
 Oy 961 gtccagactcaactctctggtgacacatctgctcgagagaacacgacacacacggttcgt 1020
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 963 gtccagactcaactctctggtgacacatctgctcgagagaacacgacacacacggttcgt 1022

[illegible]

RESULT	3
AA006786	
ID	AA006786 standard; DNA; 2337 BP.
XX	
XX	
AC	AA006786;
XX	
XX	
DT	06-MAR-1991 (first entry)
XX	
DE	Taka-amylase A gene.
XX	
KW	Alcohol; ss.
XX	
OS	Aspergillus oryzae RIB 40.
XX	
PN	JP02268685-A.
XX	
PD	02-NOV-1990.
XX	
PF	07-APR-1989; 89JP-0086787.
XX	
PR	07-APR-1989; 89JP-0086787.
XX	
PA	(JOZO-) JOZO SHIGEN KENKYUS.
XX	

DR	WPI, 1990-372009/50.
XX	
PT	New gene, vector and transformant of taka amylase A - isolated from <i>Aspergillus oryzae</i> , used to produce alcohol.
PT	
XX	
PS	Claim 1; Fig 1; 6pp; Japanese.
XX	
CC	The DNA was sequenced from a clone isolated from a <i>A. oryzae</i> RIB 40 genomic library. It can be used to express recombinant
CC	taka-amylase useful in the prodn. of alcoholic drinks and alcohols.
CC	See also AAQ06784 and AAQ06785.
XX	
SQ	Sequence 2337 BP; 630 A; 600 C; 533 G; 574 T; 0 other;

Query Match	61.3%;	Score 1063.8;	DB 11;	Length 2337;
Best Local Similarity	75.3%;	Pred. No. 1.8e-237;		
Matches 1700; Conservative	0;	Mismatches 7;	Indels 552;	Gaps 8;

[illegible]

D	b	848	cggttagctcaftgcgataatgaagctgttttaaacgcttccagcttccaaagactacttccacc	907
Q	y	556	cgcttcgtttcattcaaaaactatgaagaatcagactcaagctttagagatttgcgtcctaagag	615
D	b	908	cgcttcgtttcattcaaaaactatgaagaatcagactcaagctttagagatttgcgtcctaagag	967
Q	y	616	ataaacgtctcctcttgccttcgtatctctgcataccaacaaggaatgtgttcaagaatgaatgt	675
D	b	968	ataaacgtctcctcttgccttcgtatctctgcataccaacaaggaatgtgttcaagaatgaatgt	1027
Q	y	676	acgacgtggcgagatacttctgtatcgaactactcca	711
D	b	1028	acgacgtggcgagatacttctgtatcgaactactcca	1087
Q	y	712	-----ttgacggtccctcgatccgacacagta	737
D	b	1088	caacttgcgtatcgatgaataacttaagaatacgaatgaatgaagcgtccctcgatccgacacagta	1147
Q	y	738	aaacacgtccagaagaagactcttgcgccgggtlacaaacaaagccgagcgctgtactgtac	797
D	b	1148	aaacacgtccagaagaagactcttgcgccgggtlacaaacaaagccgagcgctgtactgtac	1207
Q	y	798	ggcgaggttgcttcgacgagtgatctccggtcactacacttgcgccttaccagaagaagtatgaagcg	857
D	b	1208	ggcgaggttgcttcgacgagtgatctccggtcactacacttgcgccttaccagaagaagtatgaagcg	1267
Q	y	858	gtactgtaactatcccat-----	874
D	b	1268	gtactgtaactatcccatgtatgttctctccaaacaaagactcttcttgcaagtctatctc	1327
Q	y	875	-----ttactatcactcctccaaagcgttcaagttcaactccg	912
D	b	1328	ctaacgaacacggtctaaacacgaacttactatccactcctccaaagcgttcaagttcaactccg	1387
Q	y	913	gcagcatagcgacgactcttaacatgatatacaacacggtcaaatccgacttccagactca	972
D	b	1388	gcagcatagcgacgactcttaacatgatatacaacacggtcaaatccgacttccagactca	1447
Q	y	973	caaccctggcgacacttctgtctgagaaacaacgacacccacggttccgttc-----	1021
D	b	1448	caaccctggcgacacttctgtctgagaaacaacgacacccacggttccgttcgaagtcttcc	1507
Q	y	1022	-----ttacac	1027
D	b	1508	ctttatttccgttccccaattccacaacagaacccacactaaagaagcaaaagttaac	1567
Q	y	1028	caaaagcatagcccttcgcgaagaagcttgcgagatttcatcatccctccaaagaacgagatcc	1087
D	b	1568	caaaagcatagcccttcgcgaagaagcttgcgagatttcatcatccctccaaagaacgagatcc	1627
Q	y	1088	catcatctaacgcgcgcgaagaacagacactacgcgcgcgcgaagaacgcgcgcgaacgcga	1147
D	b	1628	catcatctaacgcgcgcgaagaacagacactacgcgcgcgcgcgaagaacgcgcgcgaacgcga	1687
Q	y	1148	agcaaaccttgctcttcggtctacccgcgaacgcgaacgcgaacgcgaacgcgaacgcgaacgcga	1207
D	b	1688	agcaaaccttgctcttcggtctacccgcgaacgcgaacgcgaacgcgaacgcgaacgcgaacgcga	1747
Q	y	1208	gaacgcgaatccgcgaactatgcatatgcaaaagatacagaatgcgttgaactaca-----	1267
D	b	1748	gaacgcgaatccgcgaactatgcatatgcaaaagatacagaatgcgttgaactacaagtaag	1807
Q	y	1261	-----	1267
D	b	1808	cacaaccttaaccataccctaattgaccttcttcaagagtatctgaacacaagaagactaat	1867
Q	y	1261	-----agaacttggcccatcttacaagaacgacacaacagatcgcgcatgcgcgaag	1308
D	b	1868	cacttgcgaatacagaacttgcgcactcttacaagaacgacacaacagatcgcgcatgcgcgaag	1927
Q	y	1309	gcacagatgggttcgcagatctgtgactactcttgcacaagaaggttgccttgcgggtgactcgt	1367
D	b	1928	gcacagatgggttcgcagatctgtgactactcttgcacaagaaggttgccttgcgggtgactcgt	1987

QY	1369	ataccctctctcttgtagtggtgcgggtttacacagccgcccgaacgaattgacggaagtcatg	1428
Db	1988	ataccctctctcttgtagtggtgcgggtttacacagccgcccgaacgaattgacggaagtcatg	2047
OY	1429	gtctgacgaacctgacggtttggttcggatggaaatcgttcgttccctatggcaagtgggc	1488
Db	2048	gtctgacgaacctgacggtttggttcggatggaaatcgttcgttccctatggcaagtgggc	2107
OY	1489	tacctagggatttgatcccgactggaagtgttgccaggtagcaagaatctctagtagctcgt	1548
Db	2108	tacctagggatttgatcccgactggaagtgttgccaggtagcaagaatctctagtagctcgt	2167
OY	1549	gaagaggtgagagatatagagbgtactctctatcaatctgsgattggacagrtgatttga	1608
Db	2168	gaagaggtgagagatatagagbgtactctctatcaatctgsgattggacagrtgatttga	2227
OY	1609	gtttatgtacagtttgaggtcttgactgtcgtcatcccttactctctgatttttt	1668
Db	2228	gtttatgtacagtttgaggtcttgactgtcgtcatcccttactctctgatttttt	2287
OY	1669	cgaaacctaatgccaagacgcgtagtctatattaggaaa	1707
Db	2288	cgaaacctaatgccaagacgcgtagtctatattaggaaa	2326
RESULT 4			
ID	AAFI1252	standard; cDNA; 1389 BP.	
XX	AAFI1252;		
AC	AAFI1252;		
XX	13-MAR-2001 (first entry)		
DT	13-MAR-2001 (first entry)		
XX	Aspergillus niger EST SEQ ID NO:3775.		
DE	Aspergillus niger EST SEQ ID NO:3775.		
XX	Multiple gene expression; filamentous fungal cell; EST;		
KW	expressed sequence tag; Fusarium venenatum; Aspergillus niger;		
KW	Aspergillus oryzae; Trichoderma reesei; identification; recombination;		
KW	culture condition; environmental stress; spore morphogenesis;		
KW	metabolic pathway engineering; catalytic pathway engineering; ss.		
XX	Aspergillus niger.		
OS	Aspergillus niger.		
XX	WO2000056762-A2.		
PN	WO2000056762-A2.		
XX	28-SEP-2000.		
PD	28-SEP-2000.		
XX	22-MAR-2000; 2000WO-US07781.		
PF	22-MAR-2000; 2000WO-US07781.		
XX	22-MAR-1999; 99US-0273623.		
PR	22-MAR-1999; 99US-0273623.		
XX	(NOVO) NOVO NORDISK BIOTECH INC.		
PA	(NOVO) NOVO NORDISK AS.		
PA	(NOVO) NOVO NORDISK AS.		
XX	Berka RM, Rey MM, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;		
PI	WPI; 2000-594572/56.		
XX	WPI; 2000-594572/56.		
DR	Monitoring differential expression of genes in filamentous fungal cells		
XX	uses fluorescence-labeled nucleic acids isolated from the cells and a		
PT	substrate of expressed sequence tags -		
XX	Claim 87; Page 1694-1695; 3161pp; English.		
PS	Claim 87; Page 1694-1695; 3161pp; English.		
XX	The present invention describes a method for monitoring differential		
CC	expression of genes in a first filamentous fungal (FF) cell relative to		
CC	expression of the same genes in one or more second filamentous fungal		
CC	cells. The method uses fluorescence-labeled nucleic acids isolated from		
CC	the FF cells and a substrate of expressed sequence tags (EST). The ESTs		
CC	are used in the methods for monitoring differential expression of genes		
CC	in a first filamentous fungal (FF) cell relative to expression of the		
CC	same genes in one or more second filamentous fungal cells. Monitoring		

CC the global expression of genes from FF cells allows the production
 CC potential of the microorganisms to be improved. New genes may be
 CC discovered, possible functions of unknown open reading frames can be
 CC identified and gene copy number variation and stability can be
 CC monitored. The expression of genes can be used to study how FF cells
 CC adapt to changes in culture conditions, environmental stress, spore
 CC morphogenesis, recombination, metabolic or catabolic pathway
 CC engineering. Using ESTs provides several advantages over genomic or
 CC random cDNA clones including elimination of redundancy as one spot on an
 CC array equals one gene or open reading frame, and organisation of the
 CC microarrays based on function of the gene products to facilitate
 CC analysis of the results. AA07478 to AA11247 represents ESTs from
 CC Fusarium venenatum; AA11248 to AA11853 represents ESTs from Aspergillus
 CC niger; AA11854 to AA14878 represents ESTs from Aspergillus oryzae; and
 CC AA14879 to AA15337 represents ESTs from Trichoderma reesei, which are
 CC all specifically claimed in the present invention.

XX Sequence 1389 BP; 354 A; 384 C; 325 G; 326 T; 0 other;

Query Match 46.5%; Score 806; DB 21; Length 1389;

Best Local Similarity 100.0%; Pred. No. 1.2e-177;

Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 902 gtcaacctcgcgcagatgagcagaccttacaacatgatcaacacgcgtcaatccgactg 961
 DB 1 gtcaacctcgcgcagatgagcagaccttacaacatgatcaacacgcgtcaatccgactg 60
 OY 962 tccagactcaaacctccttgaggacattcgtcgagaaccacgacacccaggttcgcttc 1021
 DB 61 tccagactcaaacctccttgaggacattcgtcgagaaccacgacacccaggttcgcttc 120
 OY 1022 ttacacacagacatgacctcctgcgaagaacgfcgcagacatcatcctctcaacagag 1081
 DB 121 ttacacacagacatgacctcctgcgaagaacgfcgcagacatcatcctctcaacagag 180
 OY 1082 aatcccatcatctacgcgcgcgaagaacgacctaagccgcgcgaacacgaccccgagaa 1141
 DB 181 aatcccatcatctacgcgcgcgaagaacgacctaagccgcgcgaacacgaccccgagaa 240
 OY 1142 ccgcgaagaacactggtcttcgggtaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1201
 DB 241 ccgcgaagaacactggtcttcgggtaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 300
 OY 1202 ctccgcgaagcgcacatccgcgaacatgcatctagcaagaaggttcgtgacctacaa 1261
 DB 301 ctccgcgaagcgcacatccgcgaacatgcatctagcaagaaggttcgtgacctacaa 360
 OY 1262 gaactggcccatctacaagaacgacacagatcgcatgctgcgaagggcacagatggttc 1321
 DB 361 gaactggcccatctacaagaacgacacagatcgcatgctgcgaagggcacagatggttc 420
 OY 1322 gcaagatcggtgactatttgtccaagaaggtgtctcggtgattgctatccctccct 1381
 DB 421 gcaagatcggtgactatttgtccaagaaggtgtctcggtgattgctatccctccct 480
 OY 1382 gagggtgctgggtttacacgc 1441
 DB 481 gagggtgctgggtttacacgc 540
 OY 1442 gacggtgttgcgattggaatgtgccttccatgtagcaggtgggtacactaggtatt 1501
 DB 541 gacggtgttgcgattggaatgtgccttccatgtagcaggtgggtacactaggtatt 600
 OY 1502 gtaacgcgaatggaagttggcaggttagcagaatctgtatgctcgtgaaaggttgagag 1561
 DB 601 gtaacgcgaatggaagttggcaggttagcagaatctgtatgctcgtgaaaggttgagag 660
 OY 1562 tatatatgtaactgctatcaatctgcatgtgacagatgagattgagtttgatgacag 1621
 DB 661 tatatatgtaactgctatcaatctgcatgtgacagatgagattgagtttgatgacag 720
 OY 1622 ttggagtgcttactgctgtcatccctatactactcgtatgttttcgaaaccttaagc 1681

DB 721 ttggagtgcttactgctgtcatccctatactactcgtatgttttcgaaaccttaagc 780
 OY 1682 caagcagctagctatctattataggaa 1707
 DB 781 caagcagctagctatctattataggaa 806

RESULT 5

AA08712 standard; cDNA; 1956 BP.

AA08712;

05-JUN-1996 (first entry)

Aspergillus kawachii (IFO 4308) alpha-amylase.

Aspergillus kawachii; IFO 4308; alpha-amylase; alcohol;

starch saccharification; transformed Saccharomyces cerevisiae;

recombinant production; acid resistant; heat stable; ss.

Aspegillus kawachii.

JP0709979-A.

18-APR-1995.

30-SEP-1993; 93JP-0265447.

30-SEP-1993; 93JP-0265447.

(JOZO-) JOZO SHIGEN KENKYUSHO KK.

(KOKU-) KOKUZEI CHO CHOHAN.

WPI; 1995-182071/24.

Aspergillus kawachii alpha-amylase gene and transformant - for

recombinant prodn. of acid-resistant, thermostable enzyme useful for

direct saccharification of starch to produce alcoholic beverages

Claim 1; Page 5; 7pp; Japanese.

AA08712 encodes the A. kawachii (IFO 4308) alpha-amylase, it was used

in the construction of an expression vector for the transformation

of a S. cerevisiae host. The transformed S. cerevisiae produced

recombinant alpha-amylase, which was pH (2-7) resistant and heat

stable (80% at 60 deg. C for 60 mins.). The alpha-amylase allows

the direct saccharification of starch for the prodn. of alcohol, and

alcoholic beverages.

Sequence 1956 BP; 457 A; 569 C; 485 G; 445 T; 0 other;

Query Match 32.9%; Score 570; DB 16; Length 1956;

Best Local Similarity 64.1%; Pred. No. 8.8e-123;

Matches 910; Conservative 0; Mismatches 450; Indels 60; Gaps 1;

OY 121 ctggcagctggcgatcgatcgaatcattatctctctcgaacgattgtgcaagcag 180
 DB 71 ctggcagctggcgatcgatcgaatcattatctctctcgaacgattgtgcaagcag 130
 OY 181 atggc 240
 DB 131 acaatcgatcgaatcgatcgatcgaatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 190
 OY 241 aggc 300
 DB 191 aaggaattatcaacacatcgc 250
 OY 301 ccccgctacagccgc 360
 DB 251 cgcctacatcgagc 310

QY	361	gycgcagcagatataatactctctctgaaacgaaataacgaacgcagcagatgacttgaagcgc	420
Db	311	gycgcagcagagataatactctctgaaacgaaataacgaacgcagcagatgacttgaagcgc	370
QY	421	tctcttcgcgcctcaatgaagagggatctatcttatgtgcagatggtgtgttaacata	480
Db	371	tcttcgcagtgctctcaacgcgcgcgcgaatgtactctatgtgtcgaagtcgtctccataca	430
QY	481	tggcgctatgtatggagcgggtagctcaatcagctatcaagtggttttaaacggtcaagttccc	540
Db	431	tggcgctacgcaggttaacggcacaacgatatgtgattacaagctcttcgaccccaatc-----	483
QY	541	aagcgtacttccaccccgctctgtttcatctcaaaactatgaagatcagatcctaggttagg	600
Db	484	-----gtccaag	490
QY	601	atctgtgcctaggaagataaacatctctctctgcctgatctcgatacccaagaatgttg	660
Db	491	actgttgagggggtgacacccatcgtctctcttcgcagatctgataccacagaaacgcgcg	550
QY	661	tcaagaatgatgtatgcagactgggtggggtatcttggtatctgaacactccattgaagcc	720
Db	551	tgaagaacatttgtctacgattgggttagccgacctgtgataccaactacagtcgaagcc	610
QY	721	tcgcgtacgcacagatgaaacacagcttcgaagaagactctctgcccgggtgacaacaaagccg	780
Db	611	tcgcgtacgcagatgtctcgaaagatcgaaccccgagttcttccgggtcgaagaagaagag	670
QY	781	caggcgtgtactgtatctgcgcgaggtgtctcgaaagtgatccgcgcctacaactgttccctaac	840
Db	671	caggagctactctgtctcgtgtgaaggtcgaacaagcgaacccctgtctcgtactgtgccataac	730
QY	841	aagaagctcagagcggcgggtactctgaactatcccatctacatccatctcctcaagccttca	900
Db	731	aaaaatatctagatggtgttcttcaactatcccatctacatccatctcgtgaactctctcaagcctt	790
QY	901	agctaaccttcgcgcagcatgtgacgcacccctctaaacatgtatcaaacgcgttcaatccgact	960
Db	791	aatctccaaacggcagatcatgaacacccctctacaacatgataacatcgtctgcgcagcgact	850
QY	961	gtccagctcaaacactcctctgggacactctcgtcgagaagacagacaacccagcgtttcgctt	1020
Db	851	gtctcgatgcagccctctcctgtggcaacttcatgaaaacacagacacaccccgcttgcct	910
QY	1021	cttaacaccaaagacatcacccctctgcgcgaagaacgtcgcagcatctcatctcctcaagcag	1080
Db	911	ctctacacatccgactactctccaaagccaaaacgttctcagataactttctcctctccgacg	970
QY	1081	gaatcccatcatcttaacgcccgcgcgaagaacagcactacgcgcgcggaaacagaccgccgga	1140
Db	971	gcatctcccatctgtctaacgcgcgcgcgcgaagaagaacatcaatccgcgcgcgcgactgtccct	1030
QY	1141	acgcgcgaagcaaacctgtgctctgcgctacccgcgcagacagcgagcgtgtacataagttaatg	1200
Db	1031	acaacgcgcgagactgtgctatcagcgtctcagacacccctcgcgcgagctctcaacacttgatag	1090
QY	1201	cctcgcgcgaacgcgaatccgcgaactatgtccaattgaacaaagatacagagatctgtgactaca	1260
Db	1091	ccacccaacaaacgcgactccggaacactagtctatctcaagcaaacctcgcgactcatcttactica	1150
QY	1261	agaactgtgccaattacaagaagcagacaacacccacacgcgcgcatgctgcaaaagcactccgcgc	1320
Db	1151	agaaagaccccaatctacaacagcagacaacacccacacgcgcgcatgctgcaaaagcactccgcgc	1210
QY	1321	cgcagatcgttgaactatctgttccaacaagaaggtgtcttcgggtgtatgtatgcataccctctct	1380
Db	1211	cccaaatcatcacgcgtctccttccaacaagaaggtcctccgcggaagcagctacaccccttcaacc	1270
QY	1381	tgaatgtgtgcgggttacaacagccgcgcagcgaatctgacgcgaagtcatgtgctgtcagacgcg	1440
Db	1271	tcaagcgagggcggtctataacgttccgcgcgcgaaggtactctgaagcgttaacaccttgcagctgc	1330

OY	1441	tgacggttggttcgagagaagtcgccttcctacatggcagatgctgctactcagggtat	1500
Db	1331	tgacagctggaactcgacaagggagataccctcgctgcagctgcttcgcgagatcactcagaattc	1390
OY	1501	tgatccgactgagaagcttgcgcagctagcacaagatctctag	1540
Db	1391	tcctccctgcgttcgctgttgatgatagtctctgcgtttctgag	1430

RESULT 6

ID AAF11252 standard; cDNA; 1389 BP.

AC AAF11252;

DT 13-MAR-2001 (first entry)

DE Aspergillus niger EST SEQ ID NO:3775.

KW Multiple gene expression; filamentous fungal cell; EST;

KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;

KW metabolic pathway engineering; catabolic pathway engineering; ss.

OS *Aspergillus niger*.

PN W0200056762-A2.

PD 28-SEP-2000.

PF 22-MAR-2000; 2000WO-US07781.

PR 22-MAR-1999; 99US-0273623.

PA (NOVO) NOVO NORDISK BIOTECH INC.

XX
XX

XX
XX
RPT. 2000-504573/EE

XX
XX
Mod +04162 71490800

PT uses fluorescence-labeled nucleic acid substrates of increased accuracy.

XX
DC
C1338 87. Page 1604-1605. English

XX
XX
E

CC expression of genes in a first filamentous fungal (Ff) cell relative to

CC cells. The method uses fluorescence-labeled nucleic acids isolated from the ER colic and a substrate of expressed sequence tags (EST). The EST

CC are used in the methods for monitoring differential expression of CC in a filamentous fungal (FF) cell relative to expression of

CC same genes in one or more second filamentous fungal cells. Monitoring

potential of the microorganisms to be improved. New genes may be discovered possible functions of unknown open reading frames can be

CC identified and gene copy number variation and stability can be used to study how FE cells

CC adapt to changes in culture conditions, environmental stress, spore
CC morphogenesis, recombination, metabolic or catabolic pathway

CC engineering. Using ESTs provides several advantages over genomic or random DNA clones including elimination of redundancy as one spot on an

CC array equals one gene or open reading frame, and organisation of the
CC microarrays based on function of the gene products to facilitate

CC analysis of the results. AAF07478 to AAF11247 represents ESTs from

CC niger; AAF11854 to AAF14878 represents ESTs from *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from *Trichoderma reesei*, which are

CC all specifically claimed in the present invention.

QY	901	agttcaacctcgcgcatgtagcagacctctaaacattgtaacaacgctloaatccgact	960
Db	791	aaactactgattacaagttccagtgagttgactaaatgattccaagcgttgcttcagtt	850
QY	961	gtccagactcaactccttggcaccttgctcgagagacacgacacccacgcttcgctt	1020
Db	851	gttcgagttcaacttcttgtagacaacttcttgtagaaatcagataatgaagttcgcctt	910
QY	1021	cttaccaccaagcagatagccctcgcgaagaagctgcgagatcatcatcctcaacgagc	1080
Db	911	caatgaccacgagcccaaaagtttgattcttaattgctatttgcatttgcctttgggtgag	970
QY	1081	gaatcccaatcatctacgcgcggccgaagaacgaactagccggtcggaagaccccgca	1140
Db	971	gtatcctctgcatcttactatagagaagaacgaagctttagcgggaaaagtgcaccaaca	1030
QY	1141	accgcgaagcaactctgctcgcggtaccgcgacgcgacgagctgtacaaagttaattg	1200
Db	1031	acagagagcgcttggtgtgattacgcggtatacaacaagaagagtgaattacaagctcatg	1090
QY	1201	ccttcgcggaacgcgaatcccggaactatgcatcttagcaagaagatcagatctgtgactaca	1260
Db	1091	ccaaagactaattctgcgcgaagaagaacgcgcgttattatcaactcaagctatgccaactcgc	1150
QY	1261	agaacttggcccatctacaagaagacgaacgaacgacgtcgcatgcgacaaggcacagatgggt	1320
Db	1151	agcttcttgtagactctttccaatgacacatgattatgcaacaaaaggagcagcgttgctt	1210
QY	1321	cgcagatcgtgactatcctgttccacaagaagtgtcctcgggtgattcgatataccctcct	1380
Db	1211	ctg-----tttccaacaacctgtgtccagcggtctctctgtatgtgacta	1255
QY	1381	tgaagtgtgggtttacacagcgcgcgcgaacattgacgaggtcatctgtgcgcagacgcg	1440
Db	1256	tttccaacaacagttacagttcccggtgagagatttgttagaagttttgacatgacgactcg	1315
QY	1441	tgacggttggttcgcgatgagaaatgtgcctgttc	1473
Db	1316	tttagcgagcgtctgacttacaagtcttcattcc	1348
RESULT 8			
AAN70916			
ID	AAN70916	standard; DNA; 4214 BP.	
XX	AC		
XX	AA	AAN70916;	
DT	03-MAY-1991	(first entry)	
XX	DX		
DE	Sequence encoding alpha-amyase from plasmid pSf alpha 1.		
XX	KW	Amylase; ds.	
XX	OS	Saccharomyces fibuligera HUT7212.	
XX	Key	Location/Qualifiers	
FT	CDS	1531..3015	
FT		/tag= a	
XX	PN	JF62104576-A.	
XX	XX	15-MAY-1987.	
PD	31-OCT-1985;	85JP-0244092.	
PF	31-OCT-1985;	85JP-0244092.	
XX	31-OCT-1985;	85JP-0244092.	
XX	PA	(FUKU/) FUKUI S.	
XX	WPI;	1987-173694/75.	
XX	P-PSDB;	AAP70571.	

PT Amylase prodn. - comprises culturing microorganism transformed
PT with vector deoxyribonucleic acid, accumulating and collecting
PT amylase
PS Disclosure; Fig 1; 14pp; Japanese.
XX
XX The plasmid may be used to transform an E.coli expression system
CC the stable production of amylase, useful in ethanol fermentation.
CC See also AAN70917.
XX
XX Sequence 4214 BP; 1249 A; 784 C; 860 G; 1321 T; 0 other;
SQ

Query Match	21.8%;	Score 377.4;	DB 8;	Length 4214;
Best Local Similarity	54.8%;	Pred. No. 6e-78;		
Matches 778; Conservative	0;	Mismatches 626;	Indels 15;	Gaps 1;

QY	55	tgctgcgcgtgagcgtctattctctgtaacgcgttcaggctgcgcagcactgcttgcgtc	114
Db	1553	tgcttgcctcaattgctgcctctgtttatagtccaacagtgactctatatcaagaagaa	1612
QY	115	caagcgtgcgagctcgtgcgacgcgaatccatttatcttccttcacagatcgtatgcaa	174
Db	1613	ctaagtctgataaattgagagatcacagctctattatccaattgtcagtcagattgcta	1672
QY	175	ggaaggaatggtgcgcgcgactgcgtcacttgytaactgcggaatcaagaatactgtgtgaa	234
Db	1673	gaacgagatgtaatacaagtgtctccctgtaacacagaagatagacttactgtgtgct	1732
QY	235	catgycagagcatcatcgcgaacgttggactatatacaggaatggcgcttcacagcatc	294
Db	1733	cttccaagagcatcataaagaagtgtgatatcaatacaagatatggtgctttactgtctatt	1792
QY	295	ggaatcaccccgcttaacagcccaagctgtcccaagaccacgcgcatatgagatgctcaatg	354
Db	1793	ggattcttcacagtgttgaanaacattcccgataacacagcatatgtaatgtttacatg	1852
QY	355	gctacttgcagcagagatatatactctctcgaaacgaataactagcgcactgcaagtattga	414
Db	1853	gttactgtgatagaagaacatatcaaatatgaanaacttgcgtactgctgtatttga	1912
QY	415	aagcgctcctctcgcgcctcatatgagagggtatctatcttatgtgtcgaatgtrtgcta	474
Db	1913	agctcttgcacaagaataattgcacagatcgtgataatggttgaatggttgataatcgttacca	1972
QY	475	accataatggctatgataatggagcgggtagctcgaatcgtatcacagtgtgtttaaaccgltca	534
Db	1973	accattcagcgagatgtagcagttgagatagatcatcattactcagagatcacaccogtltca	2032
QY	535	gttcccaagactacttccaacccgctctgttttcaatccaactatgaagatcaagatcagg	594
Db	2033	acgaccacaaagactactccaataactgtgcttatttccaactatgataagcacaagctcagg	2092
QY	595	ttgagatgtctgcgtatagagataatacaactgtctccttgcgtatctcgaataccaagaag	654
Db	2093	ttcaaatgttgcgtggagaagtgtactcttccaatgtgcattaccagaatttggaaacggagaata	2152
QY	655	atgtgtgtcaagaatgaatgtgtaagcactgggtggtgatactgtgtatcgaactatccatgy	714
Db	2153	ggagacgttgcctcagcttccaattccttgggttaagaatttgtgttgcaattactaatgy	2212
QY	715	acggcctcctgtatcgacacagtaaaaacgctccgaagaagactcttcgcccgggtacaaca	774
Db	2213	atggttitaagaatgatatagttgctaaacatgttgacocaaagccttctcccgagatttgyta	2272
QY	775	aagcgcgagcgtgtgactgtatcgcgcgaggtgcctcgaacgtgtatccgcgctcaactgtgc	834
Db	2273	gtgcatctcgagtttactactcagtagcggcgaagtttccaaggaacccagcttatacatgyc	2332
QY	835	cctaccagaacgtcatggaacggcgtactgtgaactatccattactatatacctctccaag	894
Db	2333	catacaaatattacattccaaggttagtaattatccattcgtatccacccaacacagagat	2392

[illegible]

RESULT	9
AA077668	
ID	AA077668 standard; DNA; 1404 BP.
XX	
AC	AA077668;
XX	
DT	16-JUN-1995 (first entry)
XX	
DE	Variant alpha amylase deriv. from Saccharomycopsis fibuligera.
XX	
KW	alpha amylase; carbohydrate hydrolase; increased activity;
KW	lysine residue; enzyme centre; mass production; oligosaccharide;
KW	variant; cyclomaltohexan glucanotransferase; ds.
XX	
OS	Saccharomycopsis fibuligera.
XX	
FT	Key
FT	misc_difference 247..249
FT	/*tag=
FT	/note= "the wild type sequence TAY was mutated to AAC to give a variant enzyme"
XX	
PN	JP06253836-A.
XX	
PD	13-SEP-1994.
XX	
EP	04-MAR-1993; 93TP-0065303.
XX	
PR	04-MAR-1993; 93TP-0065303.
XX	
CA	(AGEN) AGENCY OF IND SCI & TECHNOLOGY.
XX	

DR WP1: 1994-338987/41.
DR P-PSDB; AAR63187.
XX
PT Variant carbohydrtase hydrolase(s) with increased activity -
PT consists of e.g. alpha-amylase with tyrosine residue in enzyme
PT centre, useful for mass-prodn. of oligosaccharide(s)
XX
PT
PS
XX
XX Example 1: Page 20-23; 27pp; Japanese.
CC AA077665-8 encode variant alpha amylases, composed by substituting
CC bases 247-249 of the structural gene region, with TTC, TGG, CTC or
CC AAC. These substitutions result in the 8rd amino acid residue
CC (tyrosine) of the wild type sequence being changed to phenylalanine,
CC tryptophan, leucine or asparagine respectively. The substituted
CC amino acid is present in the active site of the enzyme and confers
CC increased activity on the enzyme. The variants are useful for the
CC mass production of oligosaccharides. (see AA077665 for the variant
CC structure of a cyclomalto-dextrin glucoamyltransferase).
XX
SQ Sequence 1404 BP; 408 A; 272 C; 293 G; 431 T; 0 other;

[illegible]

```

QY 841 agaagctatgagcggtactgactacattccattactatccactcctaagcgcttca 900
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 731 aaatattatccaggggttagtaattccattccattccccaacagcagatttllta 790
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 901 agtcaactccgagcagctgagcagcctctacaacatgatacagcgtaacccgact 960
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 791 aaactactgattcaagttccagtgagttgactcaaatgatttcaagcgttgcctccg 850
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 961 gtccagactcaacactccctggcagcatctgctcgagaaacacagaccccggttgc 1020
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 851 gtccgagatccaaacttgttgacaactttagaanaatacagataatgaaggttcgt 910
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1021 cttaaccaagacatagccttcgcgaagacgctcgacgttcatatcctcaagacg 1080
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 911 caatgacagcagccaagaatttgaatttctatgtattgtatcttcttcttgggtatg 970
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1081 gaatcccatcatctacgcccgaagacagacactacgcccgcggaacagcccgca 1140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 971 gtattccgttactattactatgacagacaaagccttgagcggaataaagtgacccaaca 1030
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1141 accgcgaagcaacctgctcgggtacccgacgcagcagcgtgtacaaagttaattg 1200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1031 acagagagcctgtgtgtatccgctacaaacaaagagtgactattacaaagctcatg 1090
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1201 cctcgcggaagcaatccggaactatgcatctagcaagaatcagaggtctgtgacctaca 1260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1091 ccaaaagctaaagctgcccgaagacgcgcgttctatcaagatcagctatgcccctcgc 1150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1261 agaactgcccatactacaagaacacacaaagatcgccatgacgaagggccaaagtgtg 1320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1151 agcttcgtgtactctttccaatgaccatgcttatgacaaaaagagcgagcggttctt 1210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1321 cgcagatcggtactactctgtcccaaaaggtgtcgtggtgattcgtatccctccct 1380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1211 ctg-----tttccaacaacctgtgttccaagcgtctctctgtatgtacta 1255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1381 tgaagtgcgcggttacacagccgcgcgaatgaagtgatcattgtgcgcagacgcg 1440
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1256 ttccaacaacaggttacaggttccggtgaggttggtagaagtttgcatagcagttactg 1315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1441 tgaaggttggttcggaatggaatgtgcctgttc 1473
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1316 tttagcgagcagctctgactacaagttctatcc 1348
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 10

AA077665
AA077665 standard; DNA; 1404 BP.

AC AA077665;

DE 16-JUN-1995 (first entry)

DE Variant alpha amylase deriv. from *Saccharomyopsis fibuligera*.

XX alpha amylase; carbohydrate hydrolase; increased activity;

KW tyrosine residue; enzyme centre; mass production; oligosaccharide;

KW variant; cyclomaltoextrin glucanotransferase; ds.

OS *Saccharomyopsis fibuligera*.

XX Key Location/Qualifiers

FT misc_difference 247..249 /tag= a

FT JP06253836-A.

XX 13-SEP-1994.

PF 04-MAR-1993; 93JP-0069303.

```

XX 04-MAR-1993; 93JP-0069303.
XX (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
XX WPI: 1994-328987/41.
XX P-PSDB; AAR63184.
XX Variant carbohydrate hydrolase(s) with increased activity -
XX consists of e.g. alpha-amylase with tyrosine residue in enzyme
XX centre, useful for mass-prodn. of oligosaccharide(s)
XX Example 1; Page 13-15; 27pp; Japanese.
XX AA077665-8 encode variant alpha amylases, composed by substituting
XX bases 247-249 of the structural gene region, with TTC, TCG, CTC or
XX AAC. These substitutions result in the 83rd amino acid residue
XX (tyrosine) of the wild type sequence being changed to phenylalanine,
XX tryptophan, leucine or asparagine respectively. The substituted
XX amino acid is present in the active site of the enzyme and confers
XX increased activity on the enzyme. The variants are useful for the
XX mass production of oligosaccharides. (see AA077669 for the variant
XX structure of a cyclomaltoextrin glucanotransferase).
XX Sequence 1404 BP; 406 A; 272 C; 293 G; 433 T; 0 other;
XX
XX Query Match 21.6%; Score 375.4; DB 15; Length 1404;
XX Best Local Similarity 55.6%; Pred. No. 1.3e-77;
XX Matches 752; Conservative 0; Mismatches 586; Indels 15; Gaps 1;
XX
QY 121 ctgcgagctgcgcatcgcaatccattatctccttcacgagatcgattgcaagcag 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11 ctgataatgtgaatacacagctcattatcaaatgtgtaactgacaaatttgcgagacg 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 atgggtgcagcagctgcgactgttaatactgcgagatgaagaatactgtgtggaacatg 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 71 atggtgtacaagtgctcctcctgaacaagaagatagacttctgtgtgtcttccc 130
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 agggcatcatcgcaagttgactatataccaggaatgggttcacagccatctgataca 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 131 aaggcatcaaaagaattgtgattacatcaagaatgggttactgtcatttgaatt 190
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 ccccgcttaagcccgctgcgccagacccagacatagagatgctcctaagtgact 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 191 ctccagttgttgaataaacattccgataaacacagcatatggttatcgaagttctt 250
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 ggcagcagatatactctctgaacgaaactcagcgactcagatgactgaaagcgc 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 251 ggaatgaagacatatatacaaaatgaacttgtgactgctgatttgaagctt 310
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 tctctcgcccttcagagagggggtatcttatctttagtgcgattgtgttcacata 480
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 311 tggacaagaattgcaagacgctgatatgttgaatgtgatatcogttacacacatt 370
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 tgggtatgataggcgggtagctcagtcgatacagtggttttaaacogttcaattcc 540
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 371 acggcagtgtaggagtgagatagatcgaattcagaaggtacacccggttcaagacc 430
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 541 aagactactccacccgttctgttcatcaaaactatgaagtcagactcaggttgag 600
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 431 aaaagtacttccataactactgtcttatttcaactatgtatgacccaagcctcaggttcaaa 490
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 601 atgtgctgtaggagataaacactgtctccttggctgctcgcgataccaccaagatgtg 660
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 491 gtgtctggaagtgactctcagttcattacagaagttggaagaaggaatagagag 550
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 661 tcaagaatgaatgttacgactgggtggatcaatgtgatalcgaactaccattacagcgc 720
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 551 tggcctcagtttcaattcttgggttaagatttgttgcaattactcaattgatgtgt 610
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 721 tccgtatcgaacaggttaaacagctcagaagaagactctgcccgggttaacaaagccg 780
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 611 taagaattgtagtgcataacatcgtgacaagaagcttttcccgagattttgtagtcat 670
 QY 781 cagcgctgtaactgtatcgcggaagtgctgacggtgatacggccttactgtccctacc 840
 Db 671 ctgagatttactcagtagagcggaagtttccaagaagaccacgtatacatatgccatcc 730
 QY 841 aagaagtcatacgtgacggtgactgtaactatccattactatccactccctcaacgctca 900
 Db 731 aaattaccatccagggttagtaattatccattgtaacttccatcccaacacagagattttta 790
 QY 901 agtcaacccctccgacagacatgagagacctctacaacatgatacaacacggtcaatccgact 960
 Db 791 aaactactgtattcaagttccagtgagtgactactcaaatgattccaagctgtgtccagtt 850
 QY 961 gtccagactcaaacactccctggtgacatctgtcgagacacacagacacacacggttcgt 1020
 Db 851 gtccgataccaacttggttgacaaacttggtagaaatacagataatgaagttccgctt 910
 QY 1021 cttaacacacacagacatagccctcgccaagaacgtcgacgactatccatccctcaacgacg 1080
 Db 911 caatgacccagcgaccaaagtttgatttctaattgtctatgtcattgttcctttgggtgag 970
 QY 1081 gaatcccatcatcactacgcccgcgaagaacacgacactacgcccgcggaacacgcccgcga 1140
 Db 971 gtaatccgtcattactatgacaaagacgaagcgttgagcggaanaagtgacccaaca 1030
 QY 1141 accgcgaagaacactgctcgtcggtacccgacacgacgacgacgtctagacagttaatg 1200
 Db 1031 acagagagccctgtgtgttatccggtatacaaaagagagtgactattatacaagctcatg 1090
 QY 1201 cctcgcgaacgacatccggaactatgcatatgaagaagatacagagatcgttgacctaca 1260
 Db 1091 ccaaaactaattgtcgcaaaaacgcgcgcttatacaagactcaagctatgcccactcgc 1150
 QY 1261 agaacgtgccactctacaagaacgacaaagatcgccgtcgcaaggacagatgagtt 1320
 Db 1151 agcttctgtatcttccaatgaccatgattatgcaacaaaagaagcagcgttgttt 1210
 QY 1321 cgcagatcgtgactatctgtgccaagaagtgctcgtggtgattcgtataccctctcct 1380
 Db 1211 ctg-----tttcaacaaccttggttccacgsggtcttcttgtagtgaacta 1255
 QY 1381 tgaagtgtggtggttacacaacgcccgcgaacatgacggaagttacattggtcgcagacgg 1440
 Db 1256 ttccaacacaggttacagttccggtgaggaattgtgtagaagttttgacatgcagtaactg 1315
 QY 1441 tgaagttgttcgagtggaatgtgcctgttc 1473
 Db 1316 tttagcgcgacgtcctgacttacaagtttctatcc 1348
 RESULT 11
 AAF11314
 ID AAF11314 standard; cDNA; 483 BP.
 AC AAF11314;
 DF 13-MAR-2001 (first entry)
 DE Aspergillus niger EST SEQ ID NO:3837.
 XX
 KW Multiple gene expression; filamentous fungal cell; EST;
 KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
 KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
 KW culture condition; environmental stress; spore morphogenesis;
 KW metabolic pathway engineering; catabolic pathway engineering; ss.
 OS Aspergillus niger.
 PN WO200056762-A2.
 PD 28-SEP-2000.
 XX

PF 22-MAR-2000; 2000WO-US07781.
 XX
 PR 22-MAR-1999; 99US-0273623.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.
 PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
 XX WPI: 2000-594572/56.
 XX
 PT Monitoring differential expression of genes in filamentous fungal cells
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a
 PT substrate of expressed sequence tags -
 PS
 PS Claim 87; Page 1717; 3161pp; English.

CC The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring
 CC the global expression of genes from FF cells allows the production
 CC potential of the microorganisms to be improved. New genes may be
 CC discovered, possible functions of unknown open reading frames can be
 CC identified and gene copy number variation and stability can be
 CC monitored. The expression of genes can be used to study how FF cells
 CC adapt to changes in culture conditions, environmental stress, spore
 CC morphogenesis, recombination, metabolic or catabolic pathway
 CC engineering. Using ESTs provides several advantages over genomic or
 CC random cDNA clones including elimination of redundancy as one spot on an
 CC array equals one gene or open reading frame, and organization of the
 CC microarrays based on function of the gene products to facilitate
 CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
 CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
 CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
 CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
 CC all specifically claimed in the present invention.
 CC
 XX
 SQ Sequence 483 BP; 121 A; 99 C; 137 G; 124 T; 2 other;

Query Match 21.6%; Score 374.2; DB 21; Length 483;
 Best Local Similarity 91.0%; Pred. No. 1.8e-77;
 Matches 443; Conservative 0; Mismatches 35; Indels 9; Gaps 4;

QY 1243 caggaattcgtgacctacaagaactggccatctacaagaacgacacacagatcgcatgc 1302
 Db 1 caggaattcgtgacctacaagaactggccatctacaagaacgacacacagatcgcatgc 60
 QY 1303 gcaaggcacaagatgggtgcagatcgtgactatctgttccaagaagtggtcctgggtg 1362
 Db 61 gcaaggcacaagatgggtgcagatcgtgactatctgttccaagaagtggtcctgggtg 120
 QY 1363 attcgtaacctctccttgtagtgcgtgggttacacaacgcccgcgaacatgacggaag 1422
 Db 121 attcgtaacctctccttgtagtgcgtgggttacacaacgcccgcgaacatgacggaag 180
 QY 1423 tcaatgctgcacgacccgtgacggttggttcgagatggaatggtcgttctcattgacg 1482
 Db 181 tcaatgctgcacgacccgtgacggttggttcgagatggaatggtcgttctcattgacg 240
 QY 1483 gtgggtaactaggtatctatccgacatgagaagttgacgagtagcaagaatctgtagta 1542
 Db 241 gtgggtaactaggtatctatccgacatgagaagttgacgagtagcaagaatctgtagta 300
 QY 1543 gctcgtgaagggtggaagatataatgtagtgcacatcaactctggcatggaca -gt 1600
 Db 301 gctcgtgaagggtggaagatataatgtagtgcacatcaactctggcatggacaagtg 359

```

OY 1601 gagtttagttgtagtgcagttgagatcgtctactgtcgtcacccttactcttcca 1660
    |||
Db 360 aagtttagttgtagtgcagttgagatcgtctactgtcgtcacccttactcttcatg- 418
OY 1661 ttgttttcgcacccttaatgcacagccagctagtctctatttagaanaaaaaa 1720
    |||
Db 419 ----gttttcgaactaatgtccaa-cacgcgtagtctatataagaaagntccgattaaa 473
OY 1721 aaaaaa 1727
    |||
Db 474 aaaaaa 480

RESULT 12
AAQ77667
ID AAQ77667 standard; DNA; 1404 BP.
XX
AC AAQ77667;
XX
DE 16-JUN-1995 (first entry)
XX
DE Variant alpha amylase deriv. from Saccharomyopsis fibuligera.
XX
KW alpha amylase; carbohydrate hydrolase; increased activity;
KW tyrosine residue; enzyme centre; mass production; oligosaccharide;
KW variant; cyclomaltohextrin glucanotransferase; ds.
XX
OS Saccharomyopsis fibuligera.
XX
FH Key Location/Qualifiers
FT misc_difference 247..249
FT /tag= a
FT /note= "the wild type sequence TAY was mutated to
CTC to give a variant enzyme"

JP06253836-A.
XX
PD 13-SEP-1994.
XX
PF 04-MAR-1993; 93JP-0069303.
XX
PR 04-MAR-1993; 93JP-0069303.
XX
PA (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
XX
DR WPI: 1994-328987/41.
DR P-PSDB: AAR63186.
XX
XX Variant carbohydrate hydrolase(s) with increased activity -
PT consists of e.g. alpha-amylase with tyrosine residue in enzyme
PT centre, useful for mass-prodn. of oligosaccharide(s)
XX
PS Example 1: Page 18-20; 27pp: Japanese.
XX
XX AAQ77665-8 encode variant alpha amylases, composed by substituting
CC bases 247-249 of the structural gene region, with TTC, TGG, CTC or
CC AAC. These substitutions result in the 83rd amino acid residue
CC (tyrosine) of the wild type sequence being changed to phenylalanine,
CC tryptophan, leucine or asparagine respectively. The substituted
CC amino acid is present in the active site of the enzyme and confers
CC increased activity on the enzyme. The variants are useful for the
CC mass production of oligosaccharides. (see AAQ77669 for the variant
CC structure of a cyclomaltohextrin glucanotransferase).
XX
SQ Sequence 1404 BP; 406 A; 273 C; 293 G; 432 T; 0 other;

Query Match 21.6%; Score 373.8; DB 15; Length 1404;
Best Local Similarity 55.5%; Pred. No. 3.1e-77;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;
OY 121 ctgcgcagctgcgcagatcgaatccattattctcttcacagatgatttgcagaagcg 180
    |||

```

```

Db 11 ctgataaatgagatcacagtlcatattatccaattgtcactgacagatttgcagaacg 70
OY 181 atgggtgcgcagctctgcagcttgtaactgtgcgatacagaataactgtgtgaaatgac 240
    |||
Db 71 atggtgatacaagtgcttccttcttaacacagaagaatagacttactgtgtgtcttccc 130
OY 241 agggcatcatcacaagaattgtgataatccagaagaaatgggtcttcacgacatctgata 300
    |||
Db 131 aagcatcatcaagaagattgtgattatcaatcaagatattggtcttactgtatattgatt 190
OY 301 ccccgcttaacgccagcgtcccccagacacgcgcatatggaatgcttaccatgtctact 360
    |||
Db 191 ctccagttgttgaaaaacattcccgataacacagacagatattgttattcatttccctc 250
OY 361 ggacgacgagatatactctctgaacgaacactacagcactcaagatgactgaagagcgc 420
    |||
Db 251 ggaatgaagaacatatacaaatatgaacttggtaactgctgactgtgtgttgaagctct 310
OY 421 tctcttcggccttcatgagaggggagtgatcctatggttcgaatgtgtgtcttaaccata 480
    |||
Db 311 tggcacaagaattgcacgaltcgtgatatgttgaatggttgatatacgttaccacacalt 370
OY 481 tgggtatggttgagcgggtgagctcagttacagtggttttaacgggttcagttccc 540
    |||
Db 371 acggcagtgatgagcagtgagagatgatcagatataccagagttacacccggttcaagacc 430
OY 541 aagactactccacccgcttctgttcatcaaaaactatgaagatcagactcaggttgaag 600
    |||
Db 431 aaaaatcttcccaactactatgtcttattcaaacatgatgacacaggtcaggttcaaa 490
OY 601 attgtgtcagtagagataaacactgtctctctgctgtatctcgataccacacagatgtg 660
    |||
Db 491 gtgtcgtggaagtgtaactcttcagttgcatcacaatctgagaagaagaatgaagcag 550
OY 661 tcaagaatgaaigtgaactgtgtgagatcatgtgatcgaactacatcgaagcgc 720
    |||
Db 551 tggcctcagtttccaactcttggttaagaatttgttggcaatttactcaatttgaagtt 610
OY 721 tccgtatcgacacagtaaaacacgtccagaagagactctgcccgggtacaaacacgcg 780
    |||
Db 611 taagaattgtagtgctaaacatgttgaccaaagcttcttcccggtatttgttagtgat 670
OY 781 caagcgtgtagtctatcgcgcgaggtgtcgtcgaagtgatccggcgtacacttgcctacc 840
    |||
Db 671 ctgagttactcagtaggcgaagtttccaaggaagccacagctataacatgccacacc 730
OY 841 agaacgtcatgagcgcgtactgaactcaccattactaccactcctcaacgccttca 900
    |||
Db 731 aaattatcaatccaggggttagtataatcatcatgtactaccacacagagattttta 790
OY 901 agtcaacctccgcagcatgtagcagcctctacaacatgatcaaacacgcgtcaaatcgact 960
    |||
Db 791 aaactactgattcaagtctcagtgagtgtactcaaatgttcaagcgtgtctccagtt 850
OY 961 gtccagactcaaacactctctggcacattctgcgggaacccacgaaccccggttcctt 1020
    |||
Db 851 gtccgatacccaacttgttgcaaacacttgtgaagaatcagaataaagaaggttcgtct 910
OY 1021 cttaacccaagacataagccctcgcacaagaagcgtcgaagcttcatcatcctcaagacg 1080
    |||
Db 1081 gaatcccatcatcctacgcgcgcgaagaagacagactacgcgcgcgaagaacccgcgga 1140
    |||
Db 971 gttatccctgattactatgatacagaagaagcgttgagcgcgaagaatgacccaaca 1030
OY 1141 accgcgaagcaacctggtctcgtccgtaccacgcgacgcgaagcgtctaacagttaatg 1200
    |||
Db 1031 acagagagagccttgggttaccggtctacacaaagaagaatgactattaacagctcttg 1090
OY 1201 cctccgcgaacgcaatccgcgaactatgcatatgaagaagatgagatctgtgacttaca 1260
    |||
Db 1091 ccaagactaatgtctgcagaagaagcgcgctttatcagactcaagctatgcacactgcg 1150

```

```
OY 1261 agactgcccacatacaagacacacagatcgccatgccaaggacacagatggt 1320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1151 agcttctgtgactcttcaatagaccatgcatctgcaacaagaagcagcgttctt 1210
OY 1321 cgcagatcgtagactctgtccacaagaagtgctcgggtgaltcgttacccctct 1380
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1211 ctg-----tttcaacaacctgtctccagcgctctctctgtgacta 1255
OY 1381 tgagtgtgcggttacacagccgccaagcatctgacgaggtcaattgctgcacagcg 1440
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1256 ttccaacaaggtttacagcttcgcgggtgaggttgtagaagatttgacatgcagctacty 1315
OY 1441 tgaaggttggttcgagtggaatgtgcctgtc 1473
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1316 ttacgcgcagcctctgacttacaagttctatcc 1348

RESULT 13
AA077666
ID AA077666 standard; DNA; 1404 BP.
XX
AC AA077666;
XX
DT 16-JUN-1995 (first entry)
XX
DE Variant alpha amylase deriv. from Saccharomyopsis fibuligera.
XX
KW alpha amylase; carbohydrate hydrolase; increased activity;
KW tyrosine residue; enzyme centre; mass production; oligosaccharide;
KW variant; cyclomaltohextrin glucanotransferase; ds.
XX
OS Saccharomyopsis fibuligera.
XX
FH Key location/Qualifiers
FT misc_difference 247..249
FT /*tag=a
FT /*note="The wild type sequence TAY was mutated to
FT TGG to give a variant enzyme"
XX
PN JP06253836-A.
XX
PD 13-SEP-1994.
XX
PF 04-MAR-1993; 93JP-0069303.
XX
PR 04-MAR-1993; 93JP-0069303.
XX
PA (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
XX
DR WPI: 1994-3328987/41.
DR P-PSDB: AAR63185.
XX
PT Variant carbohydrate hydrolase(s) with increased activity -
PT consists of e.g. alpha-amylase with tyrosine residue in enzyme
PT centre, useful for mass-prodn. of oligosaccharide(s)
XX
PS Example 1: Page 15-17; 27pp; Japanese.
XX
XX AA077665-8 encode variant alpha amylases, composed by substituting
CC bases 247-249 of the structural gene region, with TMC, TGG, CMC or
CC AAC. These substitutions result in the 83rd amino acid residue
CC (tyrosine) of the wild type sequence being changed to phenylalanine,
CC tryptophan, leucine or asparagine respectively. The substituted
CC amino acid is present in the active site of the enzyme and confers
CC increased activity on the enzyme. The variants are useful for the
CC mass production of oligosaccharides. (See AA077669 for the variant
CC structure of a cyclomaltohextrin glucanotransferase).
XX
SQ Sequence 1404 BP; 406 A; 271 C; 295 G; 432 T; 0 other;
```

```
Best Local Similarity 55.5%; Pred. No. 3..1e-77;
Matches 751; Conservative 0; Mismatches 587; Indels 15; Gaps 1;

OY 121 ctgcgacgtgcgcatcgcaatccattatctccctccacgggttcgtgaagagcg 180
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 11 ctgtaaatagagatccacgtctctattataaattgccaatgacagattgcgacagcg 70
OY 181 atggttcgaagactgcgacttgttaatacttcgcgaltcagaataactgtgtgtaacatgyc 240
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 71 atggtgaataagtgcttcctcgtaacacagaatagaattactgtgtgtctctcc 130
OY 241 aggcacatcagcaagaatttgactatataccaggaatagggcttccacagccatctgataca 300
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 131 aaggcacaataaagaatttgattatcaataaagatagggcttactgcttactgtgattt 190
OY 301 ccccggttaagccagctgcgcacagacacacagccacataatgagatgcttaccatgctact 360
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 191 ctccagttgttgtaaaataatcccgataacacagacataatggtttatcataggttctgt 250
OY 361 ggcagcagagatatactctctgaacgaanaaacacagcagactgacatgtaagcgcg 420
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 251 ggaatgaagaacataatacaaatatgaataaacttgtagctgtgattgtaagttct 310
OY 421 tctcttcgccccttcattgagaggggagatattatgtgtcgtgtgtgtgtaaccata 480
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 311 tggcacaagaattgcagacgcgcgatalgttltlaatggtgagatcgtttaccacacatt 370
OY 481 tggcgtatgattgagcgggtgagtcagtcagctaccagtggtgttaaaccttcagttccc 540
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 371 acgcagagtggtgcagtgagtgagatagatcagatccacagatgtaaccccggttcaagacc 430
OY 541 aagactacttccaccgcttctgtttcatccaactaataatgaatcagactcaggttgaag 600
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 431 aaaaagacttccataactctgtcttattccaactatgatgacccaagctcaggttcaaa 490
OY 601 attgcgtcgaagataaacaactgtctcttcgtccatccatccgatacccaaggatgtg 660
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 491 gtgcctgggaagtgactctcttcagttgcatcaccagatttggaaggaagatgacgagc 550
OY 661 tcaagaattgaattgtagcagctgggtggatcattggtatcgaactactccattgacggcc 720
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 551 tggccccaagtttcaattctctgtgttaagaatttctgtgcaattactcaattgtgtgt 610
OY 721 tccgtatcgacaagaacacagctccagaagaagactctgcgcccgggtacaacaagccg 780
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 611 taagaattgtagtgcataacatgtagcaagagcttctcccggaatttgtttagtcat 670
OY 781 caagcgctgactgtatcgcgcaggtgtctgcagctgatalccgctacactgttccctacc 840
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 671 ctggaatttactcagtagagcgagatttccaagaggaaccagactatcatgcccatacc 730
OY 841 agaagctcagtagagcgagctagcaactatccattactatccactccctcaagcctca 900
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 731 aaatactacatccaggggttagtaattatccattgtaactacaacacacagatatttta 790
OY 901 agtcaacctccgcagcatgagcactctacaacatgatcaacacgcgtccaatccgact 960
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 791 aaactactgattcaagttccagtgagttgactccaatgatattcaagcgttgtctccagtt 850
OY 961 gtccagactcaaacactctgggcacatctcgttcgagaacccagacacaccagcttgcgtt 1020
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 851 gtctgatalccaacttgttgacaacttgttagaataatcagataatgaagaagttcgctt 910
OY 1021 cttaacacaagaacatagccctcgccaagaagcgttcgcagacttcatctcctaagagc 1080
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 911 caatgacacagcgcaacaagtttgatttcaatgcatattgtcccttltggtgattg 970
OY 1081 gaatcccatcatctacgcgcgcgaagaacagacatacgcgcggcgaagaccccgga 1140
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 971 gttctcctgtcatattactatgacaagaacaaagcttgagcggaaaagtatgacccaaca 1030
OY 1141 accgcgaagcaactggtcttcgggttaccgcgacagcagagctgtacaagttaatt 1200
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Db	1031	acagagagcgctctgtgtgtatccgcgtctacaacaagaagatgactattacaagctcat	1090
QY	1201	cctccggaacgcgaatccggaactatgcatatgacaagatatacaggaattcgtgaact	1260
Db	1091	ccaagtatagtctgcagaagaacgcgcgctttatccaagactcaagctatgcacatgc	1150
QY	1261	agaaactgcccactctacaagaagagacaacacgatgcctctgscgaaggscagaatggt	1320
Db	1151	agcttctctgtgactcttccaatgaccacatgctattctgcacaacaagaagccagctgtt	1210
QY	1321	cgacagatcgtgactatcttctccaagaaggtgtctcggtgaltcgtatataccctctc	1380
Db	1211	ctg-----tttccaacaacctgtgtccgcgcgtctctctctgattgacta	1255
QY	1381	tgaatgtgtgcggttacaacgcgcgcgcagaatattgacgaagtcaattgctgcacacgc	1440
Db	1256	tttccaacaacaggttacaagcttcggtggaagatttgttaagaatttgacacagcaagt	1315
QY	1441	tgacggttgattcgcgaatggaatgtaagtcctgttc	1473
Db	1316	ttacgcgcagctcgtacttacaagttctatcc	1348

RESULT	14
ID	AAQ06388
XX	AAQ06388 standard; DNA; 2289 BP.
AC	
XX	AAQ06388;
DT	02-FEB-1991 (first entry)
XX	
DE	AMY1 gene encoding alpha-amylase.
XX	
KW	Schwanniomyces yeast cells; expression cassette; alpha-amylase
KV	marker gene; regulon; signal peptide; terminator;
KX	autonomously replicating sequence; ss.
OS	Schwanniomyces occidentalis.
XX	
FH	Key
FT	Location/Qualifiers
CDS	550..2085
FT	/tag= a
FT	/product=alpha-amylase
FT	508..520
FT	/tag= b
FT	/function=transcription_initiation_site
FT	2086..2088
FT	/tag= c
FT	/function=signal
FT	2180..2184
FT	/tag= d
FT	/function=signal
FT	2231..2233
FT	/tag= e
FT	/function=signal
XX	
PN	EP394538-A.
XX	
PD	31-OCT-1990.
XX	
PF	28-APR-1989; 89EP-0107780.
XX	
PR	28-APR-1989; 89EP-0107780.
XX	
PA	(RHET) RHEIN BIOTECH GES.
XX	
PI	Hollenberg C, Strasser A;
XX	
DR	WPI: 1990-328670/44.
XX	P-PSTDB; AAR07574.
PT	Transformed Schwanniomyces yeast cells - contg. an expression
PT	cassette contg. regulon, DNA coding for foreign protein and

PT	terminator
XX	
PS	Disclosure, Fig 1B(1-6); 59pp; English.
XX	
CC	The expression cassette may contain a regulon (R), a signal peptide
CC	cassette (S), a foreign protein sequence (F) and a terminator (T)
CC	of which R and/or S and/or T are derived from the alpha-amylase gene
CC	of <i>S. occidentalis</i> .
CC	R is pref. a 1.8 Kb BglII-XhoI fragment (bases -1 to -540 of this
CC	sequence).
CC	S contains all or part of one of the sequences represented
CC	in AA006383-87.
CC	F is e.g. a cellulase, interleukin, insulin-like-growth factor,
CC	interferon etc.
CC	T is pref. all or part of the terminator of this sequence
CC	(bp 1537-1740).
CC	AA01 acts as selective marker for the vector carrying the expression
CC	cassette.
CC	See also AA006389.
XX	
SQ	Sequence 2289 BP; 693 A; 398 C; 424 G; 774 T; 0 other;

Query Match	21.3%	Score 369.2	DB 11	Length 2289
Best Local Similarity	54.6%	Pred. No. 41e-76		
Matches 759	Conservative	0	Mismatches 628	Indels 3
			Gaps	1
QY 121	ctggcgactggcgatcgcatccatccattattctctctccgcatcgatctggcaagcgg	180		
DB 665	ctgataaaaggaaagcaacacatcgattatccaacatcgctactgtatgatttgcagatcgg	724		
QY 181	atbggtcgcagcactgcgactctggaactcgcgcatcgaatacgaatacttgggtggnaatgagc	240		
DB 725	atggctcgcagcacagctgactgctgtttatggatggatgcgaactcgttgggtgactatata	784		
QY 241	aaggacatcatcgacaagcttggacatataccagggaaatgggtcttcacagccatctgtatac	300		
DB 785	aaggattatcgacaagtttgattatataccaagtatgggtttccatctgcgatctgtgact	844		
QY 301	ccccgcttaacgcccagctgccccagaccacgcgcatatggatggatgctacacatggttact	360		
DB 845	ccccagttgttgcgcaaatcccgcaacaatactgcttatgttggatgcttaccagtggttatt	904		
QY 361	ggcagcaggtatatactctccgcgaacaaactcgcgaactgcgactgcagatgacttgaagcgc	420		
DB 905	ggatgaaaataatltgatgaatctgaacactaatlttggtaacgctgatgatgttgaacaacat	964		
QY 421	tcctcttcgcccctcatgagaggggagtgatctatagtgtcgatgtgtgtgtgtactaacata	480		
DB 965	tagctatgcgaaattgcattccagaaagcgtgtattgatgtgtgcagcgtgtgttacaacatt	1024		
QY 481	tgggctatgtatgagcgggtagctcagtcgatacagtggtttaaaccggttcaagtccc	540		
DB 1025	atgcttgcgaacgagatgctgttccaagcgtatgattatcttaagttccaactccaatccaat	1084		
QY 541	aagaactctccacccgcttctgtttctatccaacactatggaaatgcagctcaggtttagg	600		
DB 1085	aattcttaactccaagattatgtttgtttacaataataatgatgacaaacaatgtttgaag	1144		
QY 601	attgcctgcctagagagatacaacgtctcccttgcctgctatccgataccacaagaagtgtg	660		
DB 1145	atgtctgggaaggatgatactcgaagctccctccctccagatttaagtaacccgaagataatgaag	1204		
QY 661	tcaagaatgaatggttaacgactgggttggatcatctgtatcgaactactccaattgaagcc	720		
DB 1205	ttataggatgattccaactctgggtgcagattttgttccaanaactatccaatcatgattgtc	1264		
QY 721	tcggtatcgacacagtataaacagctccagaagagcctctgycccgggtatacaacaagcgcg	780		
DB 1265	taagaattgtatagtgcaaacacagtgataacccgttcaataacgaagtlttgaagcgtt	1324		
QY 781	cagcgtgtactgtatgcggaggtgtctgcagcgtgtatccggcttaacactgtctccacc	840		

Db	1325	cttgggtttataacttaagtgtagtlaaglttaacaaaggaagatccaacttatacttlylccaatac	1384
QY	841	agaacgtlcatagacgcgctactgaactatccattactatccactcccaacgcttca	900
Db	1385	agaattatataaaggaggttaaccacatccattatatactataccagatataagattcttca	1444
QY	901	agtcacactccgcgcagcatatgacagactcctcaacatgatccaacacgcgtcaaatccgact	960
Db	1445	gtgatactctcgcgacttccagtgtagttaaattcaattcaatgatctcccaagttacagltacatt	1504
QY	961	gtccagactlcaacacactccctggtgcacatctgctgagaaacgcagcaaccccgatgcgtt	1020
Db	1505	gttcgcagctctcttcttggttggaactttatgaaacacatgcacaagttagattccat	1564
QY	1021	cttaacccaagaacatagccctctgcgccaagaagctgcagcatatcatcctcaacgacg	1080
Db	1565	cagttacctcagacacatcccttgattaaagatfacatggtctttaaatttgggtatg	1624
QY	1081	gaatcccatcatatcaagccgcgccaagaacagacactagccgcgcygaacgcgccccgga	1140
Db	1625	gtatcccaattatttattatggccagaagaaggtctcaatggtgttccgctccgcga	1684
QY	1141	accgcgaagcaacctggtcctcgcggtactccgcgcagacgcagagctgttacaagttaatg	1200
Db	1685	atagagaagcttattggttlaagtgatataatcacgattcagaatatacagctgaac	1744
QY	1201	ccctcgcgaaagcattccggaactatgacattgcacattgcaagaatcagattctggtacctca	1260
Db	1745	gtaaactaaatacaataagaatccaagtcattaaagaagatcttcgctatccaacttca	1804
QY	1261	agaacgtggcccatctacaaagcagacacaacagatcgcactgctgcgaagggcacaagtgtg	1320
Db	1805	aatcctcatgtgttctctcttcacgacctataatagccactagaaaggttagcgatg---	1861
QY	1321	cgcagatcgcgtacatattctgtccaacaagggtgcgttcggtatctcgatccctctcct	1380
Db	1862	ctaatacaactgattccattctttaaataatttggttcaaaagcggtccacaagatatctacg	1921
QY	1381	tgaatgtgcggtgttacaacagccgcgcagcaaatgtgacggaagtcatctgctgcacgacg	1440
Db	1922	tcagacaacacgcggtatctatctagtggtgataaagattatcgatatatttcttgcgaatccg	1981
QY	1441	tgaacggtgttcogataagaaatgtccgtctccatatatgccaggtggtgtactactggat	1500
Db	1982	tttagctgtgtgactccgcgaactatcctgtatcaaatlctggtggaatgccaagaatt	2041
QY	1501	tgtatccgcac	1510
Db	2042	acgtctccgtc	2051

	RESULT	15
AAFI1273		
ID	AAFI1273	standard; cDNA; 605 BP.
XX		
AC	AAFI1273;	
XX		
DT	13-MAR-2001	(first entry)
DE	Aspergillus niger	EST SEQ ID NO:3796.
XX		
KM	Multiple gene expression; filamentous fungal cell; EST;	
KM	expressed sequence tag; Fusarium venenatum; Aspergillus niger;	
KM	Aspergillus oryzae; Trichoderma reesei; identification; recombination;	
KM	culture condition; environmental stress; spore morphogenesis;	
KM	metabolic pathway engineering; catabolic pathway engineering; ss.	
XX		
OS	Aspergillus niger.	
XX		
PN	WO200056762-A2.	
XX		
PD	28-SEP-2000.	
XX		

PF 22-MAR-2000; 2000WO-US07781.
XX
XX 22-MAR-1999; 99US-0273623.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO NORDISK AS.
XX
XX
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX
XX WPI: 2000-594572/56.
XX
XX
XX Monitoring differential expression of genes in filamentous fungal cells
PI uses fluorescence-labeled nucleic acids isolated from the cells and a
PI substrate of expressed sequence tags -
XX
XX
PS Claim 87; Page 1702-1703; 361pp; English.

The present invention describes method for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. The method uses fluorescence-labeled nucleic acids isolated from the FF cells and a substrate of expressed sequence tags (EST). The ESTs are used in the methods for monitoring differential expression of genes in a first filamentous fungal (FF) cell relative to expression of the same genes in one or more second filamentous fungal cells. Monitoring the global expression of genes from FF cells allows the production potential of the microorganisms to be improved. New genes may be discovered, possible functions of unknown open reading frames can be identified, and gene copy number variation and stability can be monitored. The expression of genes can be used to study how FF cells adapt to changes in culture conditions, environmental stress, spore morphogenesis, recombination, metabolic or catabolic pathway engineering. Using ESTs provides several advantages over genomic or random cDNA clones including elimination of redundancy as one spot on an array equals one gene or open reading frame, and organisation of the microarrays based on function of the gene products to facilitate analysis of the results. AA07478 to AA01247 represents ESTs from *Fusarium venenatum*, AA01248 to AA011853 represents ESTs from *Aspergillus niger*, AA011854 to AA014678 represents ESTs from *Aspergillus niger*, AA014679 to AA015337 represents ESTs from *Trichoderma reesei*, which are all specifically claimed in the present invention.

Sequence 605 BP; 146 A; 147 C; 156 G; 148 T; 8 other;

Query Match	21.3%;	Score 368.8;	DB 21;	Length 605;
Best Local Similarity	97.2%;	Pred. No. 3.5e-76;		
Matches 418; Conservative	0;	Mismatches. 7;	Indels 5;	Gaps 4;

[illegible]

OY 456 atggtcgatgtgtgtgtgctaaccatattggtctatgatgagccggtagctcagtcgattac 515
|||||
Db 359 atggtcgatgtgtgtgtgcta--ccatatgctatgatgagccggtagct- agtcgattac 415
|||||
OY 516 agtgtgttcta 525
|||||
Db 416 agtgggttaa 425

Search completed: June 13, 2002, 17:29:12
Job time: 13493 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 13:45:42 ; Search time 3214.34 Seconds
(without alignments)
11288.978 Million cell updates/sec

Title: US-09-710-339-1
Perfect score: 1734
Sequence: 1 tcaatcaagctctccctc.....aaaaaaaaaaaaaaaaaa 1734

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

1	1734	100.0	1734	6	AX140291	AX140291 Sequence
2	1689.6	97.4	1721	8	ASNAMY	D10461 A. shirousa
3	1083	62.5	3366	8	ASOAMY1	X12725 Aspergillus
4	1065.4	61.4	2337	6	ASNTAA	D00434 Aspergillus
5	1063.8	61.3	2337	8	E02659	E02659 DNA sequenc
6	994	57.3	3160	8	AOAMY2	X12726 Aspergillus
7	992.4	57.2	3146	8	AOAMY3	X12727 Aspergillus
8	992.4	57.2	35244	8	AB021876	AB021876 Aspergill
9	989	57.0	2935	8	ASNTAA61	M3318 A. oryzae Ta
10	985.8	56.9	2970	8	AF139925	AF139925 Aspergill
11	979	56.5	2443	8	ANMAYB	X52756 A. niger amy
12	908.2	52.4	2520	8	ANMAYA	X52755 A. niger amy
13	660	38.1	2016	23	E09025	E09025 cDNA encodi
14	434.8	25.1	2239	8	LK030376	U30376 Lipomyces K
15	409	23.6	2143	8	SCSMA2	X73497 S. cyocenes
16	377.4	21.8	4215	6	E01174	E01174 DNA encodin
17	377.4	21.8	4223	8	SPALP1	X05791 Saccharomyc
18	377	21.7	1404	6	E03536	E03536 DNA sequenc
19	375.4	21.6	1404	6	E08057	E08057 DNA encodin
20	375.4	21.6	1404	6	E08060	E08060 DNA encodin
21	375.4	21.6	1404	6	I23833	I23833 Sequence 7
22	375.4	21.6	1404	6	I43344	I43344 Sequence 4
23	375.4	21.6	1404	6	I44693	I44693 Sequence 4
24	373.8	21.6	1404	6	E08058	E08058 DNA encodin
25	373.8	21.6	1404	6	E08059	E08059 DNA encodin
26	373.8	21.6	1404	6	I23830	I23830 Sequence 1
27	373.8	21.6	1404	6	I23831	I23831 Sequence 3
28	373.8	21.6	1404	6	I23832	I23832 Sequence 5
29	373.8	21.6	1404	6	I43341	I43341 Sequence 1
30	373.8	21.6	1404	6	I43342	I43342 Sequence 2
31	373.8	21.6	1404	6	I43343	I43343 Sequence 3
32	373.8	21.6	1404	6	I44690	I44690 Sequence 1
33	373.8	21.6	1404	6	I44691	I44691 Sequence 2
34	373.8	21.6	1404	6	I44692	I44692 Sequence 3
35	371.4	21.4	2160	8	S77586	S77586 Debaryomyce
36	371.4	21.4	2161	8	SOAMY	X62079 S. occident
37	369.2	21.3	2293	8	SOAMY1	X16040 Schwannomy
38	364.8	21.0	3784	8	S38381	S38381 AMY-alpha-a
39	349.6	20.2	2290	6	A05233	A05233 Artificial
40	349.6	20.2	2290	6	A08557	A08557 Unknown alp
41	349.6	20.2	2280	6	BD000146	BD000146 Microorga
42	349.6	20.2	2280	6	E29394	E29394 Microorgan
43	315.8	18.2	9451	8	D83541	D83541 Cryptococcu
44	238.8	13.8	9451	8	AF208225	AF208225 Emerice11
45	223.8	12.9	2851	8	AB008370	AB008370 Aspergill

ALIGNMENTS

RESULT 1
AX140291 LOCUS AX140291 1734 bp DNA linear PAT 31-MAY-2001
DEFINITION Sequence 1 from Patent W00134784.
ACCESSION AX140291
VERSION AX140291.1 GI:14280549
KEYWORDS
SOURCE
ORGANISM Aspergillus oryzae.
Aspergillus oryzae.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE
1 (bases 1 to 1734)
AUTHORS hisg Rd-Frantzen,H., Svendsen,A. and Pedersen,S.
TITLE Fungamyl-like alpha-amylase variants
JOURNAL Patent: WO 0134784-A 1 17-MAY-2001;
Novozymes A/S (DK)
FEATURES
source location/Qualifiers
1..1734
/organism="Aspergillus oryzae"
/db_xref="taxon:5062"
54..1550
/note="unnamed protein product"
/codon_start=1

QY	1271	catctacaagaacgacacacacgcatcgccatgcgcaaggcgacagatgggtgcgcatcgt	1330
Db	1261	CATCTACAAAGACGACACAAAGATGCCATCGCAAGGGCAGACAGATGGGTGCAGATCGT	1320
QY	1331	gacatctctgtccacaagaagggtgtctcggtgattctgtataccctctccttgagtgtgc	1390
Db	1321	GACTATCTTGTCCAAACAAAGGGTGGTCCGGTGGATTCCGATACCCCTCCCTGAGTGGTGC	1380
QY	1391	gggttacacaacgcccgcgcgcgcatttgacgaaggatcatgtgtcgcacgcagtcgaggttgc	1450
Db	1381	GGGTTCACACACCCGCCGACGAATTCAGCGAGAGTCTATTGGCTGCACGACCCGTGACGGTGG	1440
QY	1451	ttcgcgatgaaatgtgcctgttccctatgycaggtggcgctacctaagggtattgtaccgac	1510
Db	1441	TTCCGATGGAATGATGCGCTGTTCCATATGGCAGGTGGCGCTACCTAGGGATATTGATCCGAC	1500
QY	1511	tgagaagtgtgcaggttagcaaatctctagtagctcgtgaaaggtggaagatataatg	1570
Db	1501	TGAGAAGTTGGCAGGTACCAAGATCTGTAGTACGCTCGTGAAGGGTGGAGAGTATATGATG	1560
QY	1571	gtactgtatcatcaatcgcgcgcgcatttgacgaaggatgatttgatgtacagttgagatgc	1630
Db	1561	GTACTGCTATTCAATCTGGCATTTGGACAGTGAATTGAGTTTGATGTACAGTTGGAGTGG	1620
QY	1631	ttactgtgtcatccctctatctcttcgattgttttcgaacctaatgccaagcagcgc	1690
Db	1621	TTACTGCTGCTATCCCTTATATCTTTGATTGTTTTCGAACCTATATACCAAGCAGCAGC	1680
QY	1691	tagctctattatagaagaaaaa	1722
Db	1681	TAGTCTATATATAGAAAGAAATCAGATTATATA	1712
RESULT	3		
AOAMT1			
LOCUS	AOAMT1	3366 bp	DNA
DEFINITION	Aspergillus oryzae amy1 gene for alpha-amy1ase (BC 3.2.1.1).		linear
ACCESSION	X12725		PLN 10-APR-1993
VERSION	X12725.1	GI:2430	
KEYWORDS	amy1 gene; amylase-alpha.		
SOURCE	Aspergillus oryzae.		
ORGANISM	Aspergillus oryzae		
	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.		
REFERENCE	1 (bases 1 to 3366)		
AUTHORS	Wirsel, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-AUG-1988) Wirsel S., Inst. f. Biochemie, TH Darmstadt, Petersenstr. 22, D-6100 Darmstadt, FRG		
REFERENCE	2 (bases 1 to 3366)		
AUTHORS	Wirsel, S., Lachmund, A., Wildhardt, G., and Rutkowski, E.		
TITLE	Three alpha-amy1ase genes of Aspergillus oryzae exhibit identical intron-exon organization		
JOURNAL	Mol. Microbiol. 3 (1), 3-14 (1989)		
MEDLINE	89237897		
COMMENT	[2] also presents a cDNA seq (clone pGM8); pGM8 is lacking the first 42 bp of the exon when compared to genomic seq and includes a poly(A) tail.		
FEATURES	see X12725 - X12727 for A. oryzae amy1, amy2 and amy3 genes.		
source	Location/Qualifiers		
	1..3366		
	/organism="Aspergillus oryzae"		
	/strain="DSM63303"		
	/db_xref="taxon:5062"		
	/clone="lambda Stw7"		
	/clone_id="EMBL3"		
exon	552..782		
	/gene="amy1"		
	/number=1		
gene	552..2934		
	/gene="amy1"		
	/number=8		
prim_transcript	552..2934		
	/gene="amy1"		
	sig_peptide		
	615..677		
	/gene="amy1"		
	/join(615..782,838..876,962..1077,1147..1255,1324..1552,1611..1773,1839..1985,2051..2291,2371..2658)		
	/gene="amy1"		
	/codon_start=1		
	/product="alpha-amy1ase"		
	/protein_id="CAA31218.1"		
	/db_xref="GI:295921"		
	/db_xref="SWISS-PROT:P10529"		
	/translation="MVAWMSLFYIGQVAAAPALAAATPPADRSQSYIFLLDTRAPD GSTATCNTRADQKCGGTWGGIIDKLIDYIGMGFTAIWTPVTAQLPQTPAYGAYHG WYODDIYSLNENGTADDLKALSSALHERGMYLMDVAVANHMGYDGAASSVDSEFP FSSODYPHPCFIQNYEDQVEDECMIGDNTVSLPDDITKRDYKNEMYDWSLVSN YSIDGLRIDFVKHVKDQFMPGYNKAAGVYICIGVLDGDDPATYCYQVMNDGVLNPIY YPLLNKSTSGSMDLXNMINTYKSCPSPTSLIGTPVENHNDPRFASYNIDIALAKN VAAFTILNDISPIITYAGQEHVAGNDPANREKATWLSGFTDSELYKLISANIRNY AISKDNGEVYKMPYIKDDPTTAMRGTDGSOIVTLLSNKGASGDSYTLISGAGYT AGQDLREIVICTVTVGSDGNVPMAGGLPRVLYPYEKLAGSKICSSS"		
	/join(678..782,838..876,962..1077,1147..1255,1324..1552,1611..1773,1839..1985,2051..2291,2371..2655)		
	/gene="amy1"		
	/product="alpha-amy1ase"		
	783..837		
	/gene="amy1"		
	/number=1		
exon	838..876		
	/gene="amy1"		
	/number=2		
	877..961		
	/gene="amy1"		
	/number=2		
intron	962..1077		
	/gene="amy1"		
	/number=3		
	1078..1146		
	/gene="amy1"		
	/number=3		
exon	1147..1255		
	/gene="amy1"		
	/number=4		
	1256..1323		
	/gene="amy1"		
	/number=4		
intron	1324..1552		
	/gene="amy1"		
	/number=5		
exon	1553..1610		
	/gene="amy1"		
	/number=5		
exon	1611..1773		
	/gene="amy1"		
	/number=6		
intron	1774..1838		
	/gene="amy1"		
	/number=6		
exon	1839..1985		
	/gene="amy1"		
	/number=7		
intron	1986..2050		
	/gene="amy1"		
	/number=7		
exon	2051..2291		
	/gene="amy1"		
	/number=8		
intron	2292..2370		
	/gene="amy1"		
	/number=8		
exon	2371..2934		
	/gene="amy1"		
	/number=9		
polyA_site	2934		
	/gene="amy1"		

BASE COUNT 924 a /note="pgm8" 865 c 755 g 822 t
ORIGIN
Query Match 62.5% Score 1083; DB 8; Length 3366;
Best Local Similarity 75.8% Pred. No. 1.5e-244;
Matches 1707; Conservative 0; Mismatches 0; Indels 544; Gaps 8;
QY 1 tcaatcaagctccctctctctgaacaataaaccacagaagccttaatatgctcg 60
DB 565 TCACATCAAGCTCTCCCTCTCTGAACAAATPAACCCACAGAAAGCATTTATGATGCTCG 624
QY 61 cgttggtgctctatcttctgtacagcctcaagtcgagcagccttgctgtcaagc 120
DB 625 CGTGTGCTCTATTCTGTAGAGCCTTCAGTCGCGGACCGCTTGCTTGCTCAACGC 684
QY 121 ctgcgagctgagcagcgaatcattatctctctcaagatcgattgcaagacg 180
DB 685 CTGCGGACTGGCGATCGCAATCCATTATTTCTTCAGGATCGATTGCAAGGACGG 744
QY 181 atgggtcagacgactgagctgtatctgagtc----- 216
DB 745 ATGGGTGACGACTGCGACTGTGTAATCTGCGGATGAGGTGTTGTACCTACTAGCTT 804
QY 217 -----agaatactgtgtggaacatgcaagc 245
DB 805 TCAGAAAGAGGAAATGTAACTGACTTGATATAGAAATACGTGTGTGGACATGCGAGGC 864
QY 246 atcatcgaca----- 255
DB 865 ATCATCGACAAGGTAAATTTGCCCTTTATCAAAAAAGAGGAAAGCAGAAAGAAA 924
QY 256 -----agttgacctatctcaaggaaatgg 280
DB 925 TAAATTAAGAACTAGTCTTAACCATCATGATGTTGACTATATCCAGGGAAGGG 984
QY 281 ctcaagacatctggaatcaaccccgltacagccagctgcccagacacacgacatg 340
DB 985 CTTGACACCACTCTGAGATCACCCCGTTTACAGCCAGCTGCCCCAGACACCGCATATGG 1044
QY 341 agatgctactacatgctactctgagcagagatat----- 373
DB 1045 AGATGCCCTACCATGGCTACTGGCAGAGATATTAAGTCGATTTCTTAAATATCTACC 1104
QY 374 -----atacctctgaaacgaaa 391
DB 1105 TGTCACTTTTACATCAATATGAACCTAATGATGTTTAGATACCTCTGAAACGAAAA 1164
QY 392 ctacgacatcgagatgacttgaagcgctctctcgccctcatgagaagggagatga 451
DB 1165 CTAGCGGCACTGCAGATGACTTGAAGGGCCTCTCTCGGCCCTTATGAGAGGGGATGTA 1224
QY 452 tcttatgctgcatgtggtgtctaaacatat----- 481
DB 1225 TCTTATGGTCGATGGTGGCTTAACCAATATGTTCTGTGTCCTTTGCAACTGATTCGCG 1284
QY 482 -----gggctatgtatgagggggtgagc 503
DB 1285 GATATGGTTCATTTCAGTACTGACAAATGATATATCAAGGGCTATGATGAGGCGGATAGC 1344
QY 504 tcaagtcgaatcaagtggtttaaacggttcaagttcccaagactactccaccgcttct 563
DB 1345 TCAGTCGATTCAGTGTCTTTAAACCGTTAGTCCCAAGACTACTTCCACCCCTTCTGT 1404
QY 564 ttcattcaaaactabgaagatcagactcgaagttgaagatgtgcgtgtagagaataact 623
DB 1405 TTCATTCAAAACTATGAAAGATCAACACTCAGGTGAGATTGCTGGCTAGGAGATAAACA 1464
QY 624 gtctcctgctgatatctgatatacacaagaagatgtgtgcaagaatgatatgtagactg 683
DB 1465 GTCTCTTGCCCTGATCTCGATACCAACCAAGAGATGTGTCAAGAAATGATGGTACGACTGG 1524

QY 684 gtyggatcatgtgtatcgaactactcca----- 711
DB 1525 GTGGATTCATGTGTATCGAACTACTCGAATATTTCTCCCTCATTTACAACTTGG 1584
QY 712 -----ttagcgctctcgatcgacaaagttaaacagct 745
DB 1585 CTGATGATGATACTTACGAAATCAAGTTGACGGCCCTCCGTACGACAGATAAACACAGT 1644
QY 746 ccaagaagacttcgcccgggttacacaagaacgcgaagcgtgtactatcgagaggt 805
DB 1645 CCAGAAAGACTTGTGCGCCGGGTACACAAAGCCGAGGCGTGTCTGTATGGCGAGGT 1704
QY 806 gctcgacggttgatccggtctacactgtlccctcacaagaagtcatacgagcgtactgaa 865
DB 1705 GCTGCAAGGTGATCCGGCTTACACTGTGCTTACGAAAGCATGAGAGCGGCTGTGAA 1764
QY 866 ctatcccat----- 874
DB 1765 CTATCCCATGTATGTTCTCCCAACCATGAGCCTTTTGCAAGTCATCTCTTAACGAA 1824
QY 875 -----tctatctactcctcaagccttcaagttcaacctcggaagcatg 920
DB 1825 ACGGCTTAAACAGTTTACTATCCACTCTCTCAACGCTTCAAGTCAACCTCCGGCAGCATG 1884
QY 921 gacgacctctacacacatgaatcaaacacgltcaaatccgactgtccagactcaacactctg 980
DB 1885 GACGACCTGTACACATGATCAACACCGTCAAAATCGACTGTGCCAGATCAACACTCCGG 1944
QY 981 ggcacattcgtcgagaagacagacaaacccaggttcgctc----- 1021
DB 1945 GGCACATTCGTGAGAAACACGACCAACCCACGGTTTCGTGAAGTCCTCCCTTTATTT 2004
QY 1022 -----ttacaccaagacaa 1035
DB 2005 TTCCGTTCCCAATTTTCCACACAGAACCCCACTTACACAGCAAGATTACACCAAGACA 2064
QY 1036 taagccttcgccaagaagcgtcgagcatctcatcctcacaagcagyaatcccatcatct 1095
DB 2065 TAGCCCTCGCAAGAAAGTGGAGCATTTGATCATCTCAACGACGGAATCCCATCATCT 2124
QY 1096 agcgcggtccaagaacagcaactacgctcgcgcggaagaccccggaacgcggaagcaact 1155
DB 2125 ACCCGGCGCAAGAAACAGCACTACGCGCGGGAACCAACCCCGCAACCGCGAAGCAACT 2184
QY 1156 ggtcttcgggtctaaccccgacagcagcagcgtgttaagaagttaattgctctcggaagc 1215
DB 2185 GGCTTCTGGGCTACCCGACCGACGAGCGACTGTACAGTTAATGCTCTCGGAAGCGAA 2244
QY 1216 tccgaaactatgcatctagaagaagatcacaggaatctgtgactaca----- 1260
DB 2245 TCCGGAACATATGCCATTACGAAGAATACAGGATTCGTGACCTAACGAAGTAAACACAACT 2304
QY 1261 ----- 1260
DB 2305 CTAAAGCATACCTTAATGGCTATCTTACAGATATCTGACACAAAGACTTAATCACTGGCA 2364
QY 1261 -----agaactgtgcccacttacaagaacgaacaagaatgtgcattgtgcaaggacaaagt 1316
DB 2365 ATACAGAACTGGCCCATCTTCAAAAGACGACACACACATGCCATGCGCAAGGGCACAGAT 2424
QY 1317 gggtcgacagatcgtagactactgttccaaacaaggtgtctcggtgtatctgataccctc 1376
DB 2425 GGGTTCGAGATGCTGACTATCTTGTCCAAACAAGGGGTGCTGGGTGATTCGATACCCCTC 2484
QY 1377 tcccttaagtggtgcggtttaacacagccgagcgaacttgaaggaagttcaattgtgtcgacg 1436
DB 2485 TCCTTGATGCTGCTGGGTATACACAGCCGGCCAGCAATTACGAGAGTCAATTGCTGACAG 2544
QY 1437 accgttaaggttggttcgatggaagaaatgtgctgttccatgtggaaggtgtggtactcagg 1496
DB 2545 ACCGTGACGGTGTGCTCGGATGGAATGTGCCGTTCATGTGACAGGTGGCTACTTAAAG 2604
QY 1497 gtaattatccgactgagaagttgtgcaggttagcaagatctgtagtagtcgtgaaggtgt 1556

```

|||||
Db 2605 GATATGATCCGACTGAGAAAGTTGGCAGGTAGCATCTAGTAGCTCGTGAAGGGTG 2664
OY 1557 gagagatattgtagtactgcattcaatcgtcgatggaagagtgattgattgag 1616
Db 2665 GAGAGATATGATGAGTACGTCTATTCATCTGCATGAGACAGTGTGATTGATG 2724
OY 1617 tacagttagatcgttactgcgtcctccttactactcttcgattgttttcgaacct 1676
Db 2725 TACAGTTGAGTGTGATCGTGTACGTGTCAATCCCTATATCTTTCGATTTTCGAACCT 2784
OY 1677 aatgccaagcagcagctagtctattataggaa 1707
Db 2785 AATGCCACAGCAGCTAGTCTATTATAGGAAA 2815

RESULT 4
ASNTAA
LOCUS ASNTAA 2337 bp DNA linear PLN 01-FEB-2000
DEFINITION Aspergillus oryzae Taka-amylose A (EC 3.2.1.1),
alpha-1,4-glucan-4-glucanohydrolase(TAA) gene, complete cds.
ACCESSION D00434
VERSION D00434.1 GI:217822
KEYWORDS Taka-amylose A; alpha-1,4-glucan-4-glucanohydrolase.
SOURCE Aspergillus oryzae (RIB40) genomic DNA, clone phi-AA107.
ORGANISM Aspergillus oryzae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
1 (bases 1 to 2337)
Tada, S., Jimura, T., Gomi, K., Takahashi, K., Hara, S. and Yoshizawa, K.
Cloning and nucleotide sequence of the genomic Taka-amylose A gene
of Aspergillus oryzae
Agric. Biol. Chem. 53, 593-599 (1989)
Compared with the amino acid sequence of mature TAA reported by
Toda et al., Trp-385 is inserted and Asp-476 is deleted.
FEATURES
location/Qualifiers
1..2337
/db_xref="taxon:5062"
/organism="Aspergillus oryzae"
TATA_signal
18..24
/note="putative"
<118..285
/exon
/note="Taka-amylose A precursor (exon 1)"
join(118..285,341..379,465..580,658..766,835..1063,
1122..1284,1350..1496,1562..1802,1882..2169)
/note="Taka-amylose A precursor"
/codon_start=1
/protein_id="BA00336.1"
/db_xref="gi:217823"
/translation="MMVAMWSLFLYGLVAVAPALAAPADMRSSQSIYELLTRPARD
GSTATCMTADQKCGTAVGOGIDIKLDYIOGMFTAIWTPVLAQLOPOTAVGAYHG
YWOODIYSLNENYGTADDLKALSSALHREGMYLAWDVANMAGDGSVDYVFEP
FSSODYRHPREFIONYPDQYOVKNCMDNVSLLDDTTRDYVKNMYWGVSLVN
YSIDGLRIDYKVKOKDFWPKNAAGVYCIQEVLDVDPATCTCYQVNMGVLPYI
YPLNAFSTSGSMDLLNMINTVKSDDPSTLGTVEVENDNRFASYNDDLAN
VAAFLINDGIPILYAGOECHYAGNDPANREAFMTLWGPDSLEYKLISANIRY
AISKDTGEVTKNMPYIKDFTIAMRGDSGSLVTLISNKGASGDSYTLISAGIT
AGQQLTEVICTYTVGSDGNVPYPMAGLPRLVLPFKLAGSICSSS"
121..180
sig_peptide
mat_peptide
join(181..285,341..379,465..580,658..766,835..1063,
1122..1284,1350..1496,1562..1802,1882..2166)
/product="Taka-amylose A mature peptide"
286..340
/intron
/note="Taka-amylose A intron 1"
326..331
/misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
341..379
/exon
/note="Taka-amylose A precursor (exon 2)"
380..464
/intron
/note="Taka-amylose A intron 2"
441..447
/misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
465..580

```

```

/note="Taka-amylose A precursor (exon 3)"
581..657
/intron
/note="Taka-amylose A intron 3"
639..644
/misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
658..766
/exon
/note="Taka-amylose A precursor (exon 4)"
767..834
/intron
/note="Taka-amylose A intron 4"
814..819
/misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
835..1063
/exon
/note="Taka-amylose A precursor (exon 5)"
1064..1121
/intron
/note="Taka-amylose A intron 5"
1095..1101
/misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
1108..1113
/misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
1122..1284
/exon
/note="Taka-amylose A precursor (exon 6)"
1285..1349
/intron
/note="Taka-amylose A intron 6"
1331..1337
/misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
1350..1496
/exon
/note="Taka-amylose A precursor (exon 7)"
1497..1561
/intron
/note="Taka-amylose A intron 7"
1545..1551
/misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
1562..1802
/exon
/note="Taka-amylose A precursor (exon 8)"
1803..1881
/intron
/note="Taka-amylose A intron 8"
1862..1867
/misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
1869..1874
/misc_feature
/note="consensus sequence PuCTPuAC of fungal IVS"
1882..>2169
/exon
/note="Taka-amylose A precursor (exon 9)"

BASE COUNT 630 a 598 c 535 g 574 t
ORIGIN
252 bp upstream of SalI site.

Query Match 61.4%; Score 1065.4; DB 8; Length 2337;
Best Local Similarity 75.3%; Pred. No. 2.1e-240;
Matches 1701; Conservative 0; Mismatches 6; Indels 552; Gaps 8;
OY 1 tcacatcaagctccctctctgaacaataaaccacagaaggaattatgagtcg 60
Db 68 TCACATCAAGCTCCCTCTCTGACAAATRAAACCCACAGAAAGCATTTATATGTCG 127
OY 61 cgtgtgtctctattctgttaagccttcaggctcgcgcagcctgttgcgtcaagc 120
Db 128 CGTGTGTCTCTATTTCTGACGAGCCTTCAGGTCGCGACCTGCTTGGTGAACCG 187
OY 121 ctgcgcactggcgatcgcaatcaattattcctcttcacgactgatttgaagcagg 180
Db 188 CTGCGGACTGGCGATCCCAATTCATTATTCCTTTCACGATGATTTGCAAGACGG 247
OY 181 atggtctgaagactgcgactgttaactcgcgactc----- 216
Db 248 ATGGGTGACGACTGCGACTTGTAAACTCGGATCAGTGCTGTGTACTACTAGCTT 307
OY 217 -----agaataactgltgltgaacatgcaagcagcagc 245
Db 308 TCAGAAAGAGGAATGTAACTGACTTGATATAGAAATACGTGTGTGAACATGCGAGGGC 367
OY 246 atcactgaca----- 255
Db 368 ATCATCGACAAGTAATGTCCTCCCTTTATCAAAAAAAGAGAAAGAGCAAGAAAAA 427

```


QY 256 ----- agtggactatataccaggaatgg 280
Db 428 TAAATAAAGAACTAGTCTTACCATCATCAGATGTTGACATATATCCAGGGAATGG 487
QY 281 cttcaagccctctgatacaaccccggttaacagccagcttgccecaagaccagcatatgg 340
Db 488 CTTCAACGCAATCTGATACACCCCGTTACAGCCCAAGCTGCCCCAGACACCGCATATGG 547
QY 341 agatgctcaatagctactgcaagcagatat----- 373
Db 548 AATATGCTCAATCTGCTACTGTGCACAGATATGTAAGATGCTATTTCTTTGGAAATA 607
QY 374 ----- atactctg 383
Db 608 TCTACCTGATCATATTTCTACATCATATGAACTAATGATGTTAGTATGACTCTCTG 667
QY 384 aacgaatactcagcagcagatgacttgaagcgctcttcggccctcagaaggg 443
Db 668 AAGGAAACTAGGAGCTGAGATGACTTGAAGGCGCTCTCTGCGCCCTTCATGAGAG 727
QY 444 gggatgatactatgctgactggttgtaacatat----- 481
Db 728 GGGATGATCTTATGATGATGATGCTGCTAATGATGCTTCGCTTTGCACTG 787
QY 482 ----- gggctatgtag 495
Db 788 ACTTCGCGAATGATGTTCTATTTCACTACTGACAAATGAGTAATATACGAGGCTATGAG 847
QY 496 cggtagctcagtgacttaacagtggttaacogttagtcccaactactccacc 555
Db 848 CGGGTAGCTCATGCTATACAGTGTGTTAAACGTTCACTCCCAACACTACTCCACC 907
QY 556 cgttctgttcaatcaaaactatgaagatcagactcaggttgaagatgctgagtag 615
Db 908 CTTCTGTTTCAATCAAACTATGAGATCAAGACTCAGGTGAGATGTTGGCTAGAG 967
QY 616 ataacatgctcctctgctgactgatactgataccacaagatggtgcaagaatgagt 675
Db 968 ATAAACGTGTCTCTGTTGATCTCGATACCCCAAGAGTGGTCAAGATGATGAT 1027
QY 676 acgactggtggaatcattggtatcgaaactactca----- 711
Db 1028 ACGACTGGGTGGATCATTTGGATCGAACCTCAAGTAATATTTCTCCCTCATTTCTA 1087
QY 712 ----- ttgagcgctccgfatcgacacagta 737
Db 1088 CAACTTGGCTGATGATGATCTTACGAATCAGTTGACGGCTCCGATGACAGTA 1147
QY 738 aaacgcgtccgaaggaactctgcccgggtacaacaagccgagcggtgactatc 797
Db 1148 AAACACGTCCAGAAAGACTTGTGGCCCGGTAAACAAGCCGAGGGCTGTACTGTATC 1207
QY 798 ggcgaggtgctcgaaggtgataccgactgctccctacagaagctatgaagcgc 857
Db 1208 GCGGAGGTGCTCGAGTGTATCCGGCTTACACTTGTCTTACCAAGAGTATGAGGCG 1267
QY 858 gtactgaatataccat----- 874
Db 1268 GTACTGACATATCCATGATGTCTCTCAACCATGAGCTTCTGCAATGTCATCTC 1327
QY 875 ----- tactatcaactcccaacgcttcaatgaactccg 912
Db 1328 CTAAAGAAACGGCTAAACCAAGTACTATCCACTCTCAAGCCCTTCAAGTCAACCTCCG 1387
QY 913 gaacgatgagacacttacaacatgatacacacgcttaaatcgcagctgtaacactcaa 972
Db 1388 GCAGATGAGAGACTCTTACAAATGATGATCAACACCGTCAATTCGATTTGTCACAGTCAA 1447
QY 973 caactctggagacattctgcagaacacgaacacacggttcgcttc----- 1021
Db 1448 CACTCTGGGACATCTGCTGAGAAACGACAAACCCAGGTGCTGCTGTAAGTCTTCC 1507
QY 1022 ----- ttaaac 1027

Db 1508 CTTTATTTCCGTTCCCAATTTCCACAGAAACCCCACTTAACAGAGCAAGTTACAC 1567
QY 1028 caaagacatagccctgcgcaagaagctgcagcatlcatatccataagaagaaatccc 1087
Db 1568 CAAGCATATGACCTTCCCAAGAACGTCGAGCATTCATATCTTCAAGACCGAATCCC 1627
QY 1088 catcatcagccggccaagaacagacactacgcgcgcggaagaagcccggaacgcga 1147
Db 1628 CATCATATGAGCCGCGCAAAACAGACTACGCCGGGGAAGACCCCGCAACCGCA 1687
QY 1148 agcaactggtctcgggtcctaccgacagcagcagctgtgtaaaagttaattcctccg 1207
Db 1688 AGCAACTGGCTCTCGGCGTACCCGACCGACGAGCTGTACAAATTAATTCCTCCG 1747
QY 1208 gaacgcataccggaactatgcatatagcaagaataagatctgtactaca----- 1260
Db 1748 GAACGATCCGGAATCATATGCAAAAGATACAGGATTCGTGACCTACAAAGTAA 1807
QY 1261 ----- 1260
Db 1808 CACAACCTCTAAGCATACCTAATGGCTATCTTCAAGATATCTGACACAAGACTAAT 1867
QY 1261 ----- agaactggcccatctacaagaacagacaacagatcgcagcga 1308
Db 1868 CACTGCAATACAGAACTGCGCCCATCTACAAAGACGACAAAGATGCGCATGCGCAAG 1927
QY 1309 gaacagatggtctgcagatcgtgactatctctgcacaagaaggtgcttgggtatcgt 1368
Db 1928 GCACAGATGGGTGCGACATGTGCTATCTTGTCAACAAAGGTGCTTGGGTGATTCGT 1987
QY 1369 ataccctcctctgaagtggtgcggtttacacagcgcgcgcgaactggaagagatcgt 1428
Db 1988 ATACCTCTCTTGAAGTGGTGGGTTACACACCGCCAGCAATGAGAGATATTG 2047
QY 1429 gctgcagcagctgacggttgcgttcggaatgaatgctgcttccatagcaagtgcg 1488
Db 2048 GCTGCAAGACCGTAGCGTGGTTCGATGGAATGTGCTTCTATGCGAGTGGGC 2107
QY 1489 taactaagatgtaactcgaactgaagaagtgcgaagtgcgaactctgaatgactcgt 1548
Db 2108 TACCTAGGATATGTATCCACTGAGAAAGTGGAGGTGACAGATCTGTAGAGCTGT 2167
QY 1549 gaaggtgagagatatagtatgtaactgatactcaactcgcgcaatggaagatgattga 1608
Db 2168 GAAGGTGAGAGATATATGATGTGCTACTGATTCATCTGGCAATGGACAGATTTGA 2227
QY 2228 GTTGTATGTACAGTGTGAGTGTACTGCTGTCATCCCTTATACCTTCGATGTTT 2287
QY 1669 cgaacctaatgccaagcagcgaactgactatlatagaa 1707
Db 2288 CGAACCTTAATGCCAAGCAGCTAGTCTATTATGAGAA 2326

RESULT 5
E02659
LOCUS 2337 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA sequence of takaamylase A gene.
ACCESSION E02659
VERSION E02659.1 GI:2170887
KEYWORDS JP 1990268685-A/1.
SOURCE Aspergillus oryzae.
ORGANISM Aspergillus oryzae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
1 (bases 1 to 2337)
AUTHORS Takahashi, K., Iimura, M., Gomi, K., Hara, M., Yoshizawa, K., Tada, S.
and Tamura, G.
TITLE NOVEL GENE AND VECTOR, TRANSFORMING SUBSTANCE USING SAME GENE AND
VECTOR AND USE THEREOF
JOURNAL Patent: JP 1990268685-A 1 02-NOV-1990;

TAX ADM AGENCY, JOZO SHIGEN KENKYUSHO: KK
 OS Aspergillus oryzae
 PN JP 1990268685-A/1
 PD 02-NOV-1990
 PI 07-APR-1989 JP 1989086787
 PI TAKAHASHI KOJIRO, IIMURA MINORU, GOMI KATSUYA, HARA MASAMICHI,
 PI YOSHIZAWA KIYOSHI, TADA SETSUZO, TAMURA GAKUZO PC
 C12N15/56, C12G3/02, C12G3/02, C12N1/15, C12N9/30, (C12N15/56, PC
 C12R1.69)
 PC (C12N1/15, C12R1.69), (C12N9/30, C12N1.69):
 CC strandedness: Single;
 CC topology: Linear;
 CC hypothetical: No;
 CC *Source: strain-RIB40:
 FH key Location/Qualifiers
 FT
 FT CDS 1<..<2337
 /Product='lakaamyase A'.
 /organism='Aspergillus oryzae'
 /db_xref='taxon:5062'
 BASE COUNT 630 a 599 c 534 g 574 t
 ORIGIN

Query Match 61.3%; Score 1063.8; DB 6; Length 2337;
 Best Local Similarity 75.3%; Pred. No. 4.9e-240;
 Matches 1700; Conservative 0; Mismatches 7; Indels 552; Gaps 8;

QY 1 taaactcaagctctccctctctgacaaataaaccacagagcattatgtagtgcg 60
 Db 68 TCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCACAGAGCATTATGAGTGCG 127
 QY 61 cgtgtgtgtctatcttctgtaagccttcagtcgcgcacgtcttggctgcaagc 120
 Db 128 CGTGTGTGTCTATTTCGTGACGGCTTCAGGTGCGGCACTCTTGGCTGCAGACGC 187
 QY 121 ctggcgagctgcgacatcgaatcttatttctctcgaagatgattgcaagaagcg 180
 Db 188 CTGCGGAGTGGCGATGCGATCCATTATTTCTCTCAAGATCGATTGCGAGACGG 247
 QY 181 atgggtcgacgactgcgactgtaactgcygalc----- 216
 Db 248 ATGGGTGCGACGATGCGATTTATTAATGCGGATCAGGTGCTTACTACTAGCTT 307
 QY 217 -----agaaatactgtgtggaacatgycagagcg 245
 Db 308 TCAGAAAGAGGATGTAAACTGACTTGATATGAAATACTGTGTGGAACATGGCAGGCG 367
 QY 246 atcatcgaca----- 255
 Db 368 ATCATGACAGCAAGCTAAATTCGCCCTTTATCAAAAAAGAAAGAAAGCAGAGAAAA 427
 QY 256 -----agttggaactatccaggggaatgg 280
 Db 428 TAAATTAATAAAGACGTCTGTCTTAACCATCAATAGTTGGACCTATACAGGGAATGG 487
 QY 281 ctccacagcaatctggatcaaccccggttaacagccagctgccccagacacgcgcatatgg 340
 Db 488 CTTCAACAGCCATCTGATACCCCGTTACACGCGTACGCCACGACCGCATATAGG 547
 QY 341 agatgcctacgaatgctactgcaagaagatat----- 373
 Db 548 AGATGCTTACCATGCTACTGGCAGCAGATATGTAAAGCATGTATTTCCTTTGAATA 607
 QY 374 -----atactcttg 383
 Db 608 TCACTCTCATATTCTTACATCATATATGAACATACTGATGATTTTATGATACCTCTCG 667
 QY 384 aacgaaactcaagcagctcagatgaactggaagcgctctctcgcccttcatagagag 443
 Db 668 AACGAAACTACGGCAGCTGACATGACTTTGAAGGCGCTCTTTCGCGCTTCATAGAGG 727

QY 444 gggatlatctatggtcgatgtgtgtgctaaccatat----- 481
 Db 728 GGGATGATCTTATGATCGATGCGTGGCTAACATATATGTCGTGCTTTGCAACG 787
 QY 482 -----gggcatatggag 495
 Db 788 ACTTCGGGATGATGTTTCATTTCAGTACTGACAAATGATGAATATATACAGGCGCTATGATGAG 847
 QY 496 cgggtgtcgtcagtcgattacagtggtttaaaccgttcagttcccaagctcttccacc 555
 Db 848 CCGGTAGCTCAGTCGATTCAGTACGATGTTTAAACCGTTCAAGTCCCAAGACTCTCCAC 907
 QY 556 cgttcgattcattcaaaactaatgaatcaagaactgaagttgaagattgtgtgcaagag 615
 Db 908 CGTTCTGTTTCATTCAAAACTATGAAGATCAAGACTAGGTTGAGTATTTGTTGGCTAGAGG 967
 QY 616 ataacactgtctctctgctcgtatcgtatatacccaagagatgtgtcaagaatgaatggt 675
 Db 968 ATAACTGTCCTCTGCTTGTGATCTCGATACCAAGAGATGTCAGAAATGATGAGT 1027
 QY 676 acgactggtgtggatcaatggtatgtaactactcca----- 711
 Db 1028 ACGACTGGGTGGATGATGTTGATCGAATCTACAGTAAGATATTTCCCTCATTTCA 1087
 QY 712 -----ttgacggctccgtaatgacagatga 737
 Db 1088 CAACCTTGCTGATTCGATGATGATCTTACGAATAATGTTGACGGCTCCGTTATCGACACAGTA 1147
 QY 738 aaacacgtccagaagagactctgcccgggttacaacaagcgcagcgtagtactglatc 797
 Db 1148 AAACACGTCCAGAAAGGACTCTGCGCGGTACAAACAAAGCCGACAGCGCTGATCTATTC 1207
 QY 798 ggcgaggtgtctgacggtgtgtccggttaactgttccctacagaagcttaatgagcg 857
 Db 1208 GCGGAGGTGTGCAAGCTGTGTCGGCTTACACTTGTCCCTACAGAAAGCTCAGAGCGCG 1267
 QY 858 gtaactgaactccat----- 874
 Db 1268 GTACTGAACATATCCCATGATGTTGCTTCCCAACCATGAGCGCTTCTGCAAGTCTCATTC 1327
 QY 875 -----ttactatcactcctcaagcctcaagtaactccg 912
 Db 1328 CTAAGAAAGCGGTAAACACAGTACTATCCATCTCCCAACGCTCAAGTCAACCTCCG 1387
 QY 913 gcagcatggaagcctcttaaacatgatacaacccgttaactcagacttccagactcaa 972
 Db 1388 GCAGCATGGAAGCCTCTTACACATGATCAACACCGTCAATCGATGTCAGACTCA 1447
 QY 973 caactcctgggcaactcgtcgaagaacacgaacacccaggttcgctc----- 1021
 Db 1448 CACTCTGTGGGCACTTGTGTGAAGACCGACCAACCCAGCGTTGCTGTGTAGTCTTCC 1507
 QY 1022 -----ttaac 1027
 Db 1508 CTTTATTTTCTGCTCCCAATTTCCACACAGAACCCCACTTAACAAGAGCAAGTTACAC 1567
 QY 1028 caagacatagcctctgcgaagaagctcgagatcatcatctctaagaaggaatcc 1087
 Db 1568 CAAGGATATAGCCCTCGCAGAGAGCTGCGCATATTCATATCTTAACAGCGAATGCC 1627
 QY 1088 catcatctacgcgycgcaagaagacactacgcgcgycggaagaccccgcaagccgca 1147
 Db 1628 CATCATCTACGCGCGCCCAAGACAGCATCTAGCGCGGGAAGACCCCGCAACCGCA 1687
 QY 1148 agcaactggtctcgtggttaccgcgacagcagcgtgtataagtaattgtctccgc 1207
 Db 1688 AGCAACCTGTGCTCCGCTACCCGACCGACGACGAGCTCTTAAGTAAATTCCTCCGC 1747
 QY 1208 gaacgcaatccggaactatgcaatgcaagaatgaagattggttactacta----- 1260
 Db 1748 GAACGCAATCCGGAATCTATGCCATTACCAAGATACAGGATGCTGATCTTACAAAGTAA 1807

Oy	1261	-----	1260	exon	552..782	/gene="amy2" /number=1
Db	1808	CACAACCTTAAGCATACCTTAATGGCCCTATCTTCAGATATCTGCACAAGAGCTAAT	1867	gene	552..2658	/gene="amy2" /number=1
Oy	1261	-----	1308	CDS	join(615..782,838..876,962..1077,1147..1255,1324..1552,1611..1773,1839..1985,2051..2291,2371..2658)	/gene="amy2"
Db	1868	CACCTGGCATACAGAACTGGCCCATCTTCAAAAGACACACACGATCGCCATGGCCANAG	1927			/gene="amy2"
Oy	1309	gcacagatgggtgcagatcgtgtaactatctgtccaagaaggtgcttcgagtgatcgt	1368			/product="alpha-pre-amy1ase"
Db	1928	GCACAGATGGGTCGAGATCGTACTATCTGTCCAAACAGAGGTGCTTCGGGTGATTCGT	1987			/codon_start=1
Oy	1369	ataccctctccttggtgtgtgcgggttaacacagccgcgcgaattgaaggatcattg	1428			/protein_id="CAA31219.1"
Db	1988	ATACCTCTCTCTTGAGTGGTGGGGTTACACAGCCGGCCACCAATTGACGAGGTCATTTG	2047			/db_xref="GI:1228077"
Oy	1429	gctgcacagccgtgacggttggttcgagtgaaatgtgctcttccttaagcaggtgagc	1488			/db_xref="SWISS-PROT:P10529"
Db	2048	GCTGCACACCGCTGACGGTGGGTTGGATGGAATGTGCTTCTCTATGGCAGGTGGCC	2107			/translacion="MMVAMWSLFLVGLQVAAAPALATPADMRQSIVFELTDREARTD
Oy	1489	taactaagggtatgtatccgactggaagttgacaggtagaagatctgtagtgcgt	1548			GSTATACNTADQKCYGCTWOGIIDKLDYIQCGGFPATMTPTATLPPTAYGAYHG
Db	2108	TACTAGAGGATATTGATCCGACTGAGAGTTGGCAGGTAGCAAGATCTGTAGTGCCTCGT	2167	sig_peptide		WMOODIYSLNENYGTADDLKALSALHBERGMILAVDVVANHMGYGGASVYSEYK
Oy	1549	gaagggtggaagatataatgagtgtactgtcttaacttcgaatcgtgcatggaagtttga	1608	mat_peptide		FSOQYFHPCFTYQVNEQTOVEDQCMIGDNTYSIDPDPTKDYKNEBYSYVSN
Db	2168	GAAGGCTGGAAGATATGATGATGTACTCTATTCATATTCGCAATTTGGACAGTGAATTTGA	2227			YSIDGIRITVTYKHYOKDMPGYNKAAGYCIIEVLIDGPATTCPIQNMDEVLTNPIT
Oy	1609	gtttatgataagttgagtcgttactgtctgcatcccttacccttactcttgattgtttt	1668	intron		YPLLNKFKSTSGSMDDLYNMIINTVRSQDPSDTLLGTETVENHNDNPRFASVTNDIALAKN
Db	2228	GTTTGATGTACAGTTGGAGTGGTACTGCTGTCTCATCCCTTATACTCTTCGATTGTTTTT	2287			VAAFIILNDGIPRIYAGQEQHYAGNDPANREATWLSQYPTDSLEYKLIASANIRNY
Oy	1669	cgaacctaatgccaagcagcagctagtctatataggaatg	1707	exon		ATSKDTEFTYKKNPFIYKDDPTIARRKGTGDSQIVYIILSNKGAAGSDSYTLISGAGYT
Db	2288	CGAACCTTAATGCCAACGACCGCTAGTCTATTATAGGAAA	2326	intron		ASQOQLEVICCTTVTVGSDGNVPEVMAGLPRVLVPTERKLAGSKICSSS"
RESULT	6					615..677
LOCUS	AOAMY2		3160 bp	linear		/gene="amy2"
DEFINITION	Aspergillus oryzae amy2 gene for alpha-amy1ase (EC 3.2.1.1).			exon	join(678..782,838..876,962..1077,1147..1255,1324..1552,1611..1773,1839..1985,2051..2291,2371..2655)	/number=8
ACCESSION	X12726			intron		/gene="amy2"
VERSION	X12726.1	GI:2442				/product="alpha-amy1ase"
KEYWORDS	amy2 gene; amylase-alpha.			exon		783..837
SOURCE	Aspergillus oryzae.			intron		/gene="amy2"
ORGANISM	Aspergillus oryzae					/number=1
	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eumortales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			exon		/gene="amy2"
REFERENCE	Wirsal,S.			intron		838..876
AUTHORS	Direct Submission			exon		/number=2
TITLE	Submitted (26-AUG-1988) Wirsal S., Inst. f. Biochemie, TH Darmstadt, Petersenstr. 22, D-6100 Darmstadt, FRG			intron		962..1077
REFERENCE	2 (bases 1 to 3160)			exon		/gene="amy2"
AUTHORS	Wirsal,S., Iachmund,A., Wildhardt,G. and Rutkowski,E.			intron		/number=3
TITLE	Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon organization			exon		1078..1146
JOURNAL	Mol. Microbiol. 3 (1), 3-14 (1989)			intron		/gene="amy2"
MEDLINE	89237897			exon		/number=5
COMMENT	see X12725 - X12727 for A. oryzae amy1, amy2 and amy3 genes. No amy2 transcripts were detected in Northern blot analyses; exons and introns in RT are given accordingly amy1 gene (X12725) Data kindly reviewed (28-Mar-1989) by Wirsal S.			intron		/gene="amy2"
FEATURES	Location/Qualifiers			exon		1147..1255
source	1..3160			intron		/number=4
	/organism="Aspergillus oryzae"			exon		1256..1323
	/strain="DSM63303"			intron		/gene="amy2"
	/db_xref="taxon:5062"			exon		1324..1552
	/clone="lambda Stw26."			intron		/gene="amy2"
	/clone_id="EMBL3"			exon		1553..1610
	join(552..782,838..876,962..1077,1147..1255,1324..1552,1611..1773,1839..1985,2051..2291,2371..2658)			intron		/gene="amy2"
	/gene="amy2"			exon		1611..1773
				intron		/number=5
				exon		1774..1838
				intron		/gene="amy2"
				exon		/number=6
				intron		1839..1985
				exon		/gene="amy2"
				intron		/number=7
				exon		1986..2050
				intron		/gene="amy2"
				exon		/number=7
				intron		2051..2291
				exon		/gene="amy2"
				intron		/number=8
				exon		2292..2370
				intron		/gene="amy2"
				exon		/number=8

exon 2371. .>2658
 /gene="amy2"
 /number=9
 BASE COUNT 842 a 815 c 737 g 766 t
 ORIGIN
 Query Match 57.3% Score 994: DB 8: Length 3160;
 Best Local Similarity 74.8%; Pred. No. 1.4e-223;
 Matches 1618: Conservative 0; Mismatches 0; Indels 544: Gaps 8;
 QY 1 toacatcaagctctccctctctgaacaataaccacagaaagcatltaatgaatgctg 60
 DB 565 TCACATCAAGCTCTCCCTCTCTGMAAATAAACCCACAGAAAGCATTTATGATGCTG 624
 QY 61 cgtgtgtctctatcttctgaaggcttcagtcgagccctgcttggctgcaagc 120
 DB 625 CGTGTGCTCTTATTTGTACGGCCCTTCAGGTCCGGCACCTGTGTGCTCAACGC 684
 QY 121 ctgcggaactgcgacatccatcttctcttcacgagatcttcaagagacg 180
 DB 685 CTGGGGAATGGGAGATCGCAATCCATTTATTCCTTCACGGATCGATTTCAGAGACGG 744
 QY 181 atgggtcgaagctgcgactgttaatactgcgagtc----- 216
 DB 745 ATGGGTGACGACTGCGACTTGTAATACTGCGGATCAAGTGTGTACTACTAGCTT 804
 QY 217 -----agaaatactgtgtggaagatgagcagggc 245
 DB 805 TCAGAAAGAGAGATGTAATCTGACTTGTATGAAATATCTGTGTGGAACATGCGAGGGC 864
 QY 246 atcatcgaca----- 255
 DB 865 ATCATCGACAAAGTAATTGGCCCTTTATCAAAAAAGAGAGAAAGCAGAAAGAAAA 924
 QY 256 -----agtgtactatatccagggagatggg 280
 DB 925 TAAAAATAAAGAACTAGTCTTAGTCCATACCATCATCATAGTTGGACATATATCCAGGAATGGG 984
 QY 281 ctctcagccatctgtgatacccccgtctacagccagctgcacagaccacgcatg 340
 DB 985 CTTCACAGCCATCTGGATACCCCGTTACAGCCAGCTGCCCCAGACCCGCAATATGG 1044
 QY 341 agatgcctaccatgactgactcagcagcagagatat----- 373
 DB 1045 AATGCTCTACCATGGCTACTGGCAGCAGAGATATGTAATTCATTTCTTAATATCTAC 1104
 QY 374 -----atacctctgaaagaaa 391
 DB 1105 TGTCAATCTTTTACATCAATATGAACTTAATTGATGTTTATGATTAATCTCGAAGAAAA 1164
 QY 392 ctaagggcactgagatgactgaaaggcctctctcgcaccccttcagaaaggagatgta 451
 DB 1165 CTACGGCACTGAGATGACTTGAAGGCGCTCTCTTCGGCCCTTCATGAGAGGGGATGTA 1224
 QY 452 tcttattgctcgaatgctgactaaccatat----- 481
 DB 1225 TCTTATGGTCATGTGTGTGTAACCATATGGTTGTGTGCTTTCGAACATGACTTCGCG 1284
 QY 482 -----gggactatgatgagcgaggtacg 503
 DB 1285 GATATGTTTCATTTTCAGTACTGACAATGAGTAATATCAAGGGCTATGATGAGCGGTAGC 1344
 QY 504 tcaagtcgatttaagtggtttaaacccgttcagttcccaagactacttccaccgttctgt 563
 DB 1345 TCAAGTCGATTACAGTGTGTTTAAACCGTTCACTCCCAAGACTACTTCCACCGCTTCTGT 1404
 QY 564 ttcattcaaaactgaagatcagactcaggttgaagatctgcgtgactagagataaacat 623
 DB 1405 TTCATTCAAAACTATGAGATCAGACTCAGTGTGAGATGTGTGGCTTGGAGATATACACT 1464
 QY 624 gtctcctgtcgtatcctgataccaagaagtgtgtccaagaatgaatgtaagactgtg 683

DB 1465 GTCCTCCCTGCTGANTCTGCATACCAACCAAGATGTGTCAAGAAATGAATGGTACGACTGG 1524
 QY 684 gtggatcattgtatcgaactacc----- 711
 DB 1525 GTGGATCATGTGATGATGCACTACTCTCAGTAAGATATTTCTCCCTCATTTCTCAACTTGG 1584
 QY 712 -----ttagcggctcgtatcgaacagtaaaacagt 745
 DB 1585 CTGATGATGATTAATCTACGAATAATCAATGTGACGGCTCCGTATCGACACAGTAAACAGCT 1644
 QY 746 ccagaaggaactctgcccgggttacaaagccgagagcggtactgtatcgtggagagt 805
 DB 1645 CCAAGAGGACTTCTTGCCCGGGTACAAACCCGAGCGGTACTGTATGTGGCGAGGT 1704
 QY 806 gctcagcgttgatccggtcactgtctcctaccagaacgycatggaagcgtaactgaa 865
 DB 1705 GCTCGAGGATGATCGGCTTACACTTGTCTCTACAGAAACGTCATGAGCGGCTACTGAA 1764
 QY 866 ctatcccat----- 874
 DB 1765 CTATCCCATGTATGTTCTCTCCCAACCATGAGCCTTCTGCAAGTCTCATCTCTAACGAA 1824
 QY 875 -----tctatcaccctcccaacgcttcaagtcaactccgagagatg 920
 DB 1825 ACGGCTAAAAACGATTACATCCATCCATCAAGCCTTCAAGTCAACCTCCGGCAGCATG 1884
 QY 921 gacgacctctacaactgaatcaacacgcgtcaaatccgactgtccagactcaactctg 980
 DB 1885 GACGACCTCTACAACTGATTCMAACACCGTCAATCCGATGTGTCCAGACTCAACACTCCG 1944
 QY 981 ggcacatcgtcgaagaacgaacgaacccaaggttcgttc----- 1021
 DB 1945 GGCACATTCGTGAGAAACACGACCAACCCAGCGGTGCTTGTGTAAGTCTTCCCTTTAT 2004
 QY 1022 -----ttagacaaagaca 1035
 DB 2005 TTCCGTTCCCAATTTCCACACAGAACCCCACTTAACAAAGCAAAATTAACACCAACGACA 2064
 QY 1036 taagccctcgaagaacgctgcagcaatcatcctcacaagcaggaatcccatcatct 1095
 DB 2065 TAGCCCTCGCGAAGAACGTGCGAGCATTCATCATCTTCAAGCGGAATCCCATCATCT 2124
 QY 1096 agcgcggccaagaacgaacgaactacgcgcggggaagaaagcccggaacgcggaagcaact 1155
 DB 2125 ACCCGCGCAAGAAAGACACTACGCGCGGGAACACCCCGGAACCCGGAAGCAACCT 2184
 QY 1156 ggcctcgggctaacccgacccgaagcagcgtgacaaagttaattgtcctcggaaagcaaa 1215
 DB 2185 GGCTCTCGGGCTAACCCGACCGACGAGCGAGCTGTACAAAGTTAATTGGCTCCGGAACGCA 2244
 QY 1216 tccggaactatgcatatgaagaagatacaggaatcgttgaactaca----- 1260
 DB 2245 TCCGGAACCTATGGCATTAGAAAGATACAGGANTCTGTAACATGATGAGACAAACT 2304
 QY 1261 ----- 1260
 DB 2305 CTAAAGCATACCTTAATGGCCTATCTTACAGATATCTGACACAAGAGACTTAATCACTGGCA 2364
 QY 1261 -----agaactggcccatctcaaaagacgaacagatcccatgcaagagacagat 1316
 DB 2365 ATACAGACTGGCCCATCTACAAAGACGACACACATATGCCATGCGCAAGGCGACAGAT 2424
 QY 1317 gggtcgcagatcgtgaactctgtccaaagaaggtgtcgggtgatttcgtatataccctc 1376
 DB 2425 GGGTCGAGATCGTGAATCTTGTCTCAACAAAGGGTCTTCGGGTATTCGTATATACCTTC 2484
 QY 1377 tccctgaatgtgcgggttacacagccggcagcaattgaaggaggtcaattgtgctgcag 1436
 DB 2485 TCTTTAGTGTGGGTGACAGCGGCAACAAATTGACGAGGTCAATTGGCTTCACAG 2544
 QY 1437 accgtgcaggtgtgttgagtggaatgtgcgtgtcctatlgcgaagtgagctaacctag 1496
 DB 2545 ACCGTGACGGTGTGTCGATGGAATGTGCTGTCTATGCGACAGGTGGCTACTAGG 2604

QY 1497 gattgtatccgactgagaattgacagtagtagcaatctctgtaagctcgtgaagggtg 1556
|||||
DB 2605 GTATTGTATCCGACTGAGAAAGTGGCAGGTAGCAAGATCTGTAGTACTGCTGTAAGGGTGG 2664
QY 1557 gagaatcatgatgtagctatctcaatctgcatctgagacagtgagttgagttgatg 1616
|||||
DB 2665 GAGACTATATGATGCTACTGCTATTCATCTCGCATTCGACATGGAGAGAGTTTGAGTTGATG 2724
QY 1617 ta 1618
||
DB 2725 TA 2726

RESULT 7
AOAMY3 3146 bp DNA linear PLN 10-APR-1993
LOCUS Aspergillus oryzae amy3 gene for alpha-amyase (EC 3.2.1.1).
DEFINITION X12727
ACCESSION X12727
VERSION X12727.1 GI:2454
KEYWORDS amy3 gene; amyase-alpha.
SOURCE Aspergillus oryzae.
ORGANISM Aspergillus oryzae
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
1 (bases 1 to 3146)
Wirsal,S.
Direct Submission
Submitted (26-AUG-1988) Wirsal S., Inst. f. Biochemie, TH
Darmstadt, Petersenstr. 22, D-6100 Darmstadt, FRG
2 (bases 1 to 3146)
Wirsal,S., Lachmund,A., Wildhardt,G. and Rutkowski,E.
Three alpha-amyase genes of Aspergillus oryzae exhibit identical
intron-exon organization
Mol. Microbiol. 3 (1), 3-14 (1989)
89237897
JOURNAL
MEDLINE
COMMENT see x12725 - x12727 for A oryzae amy1, amy2 and amy3 genes. amy3
transcripts were detected in Northern blot analyses [2]. Data
kindly reviewed (28-Mar-1989) by Wirsal S.
FEATURES
source location/Qualifiers
1..3146
/organism="Aspergillus oryzae"
/strain="DSM63303"
/db_xref="taxon:5062"
/clone="lambda Stw28"
/clone_lib="EMBL3"
552..782
/number=1
614..2657
/gene="amy3"
614..676
/gene="amy3"
join(614..781,837..875,961..1076,1146..1254,1323..1551,
1610..1772,1838..1984,2050..2290,2370..2657)
/gene="amy3"
/codon_start=1
/product="alpha-amyase"
/protein_id="CAA31220.1"
/db_xref="GI:295922"
/db_xref="SWISS-PROT:P10529"
/translacion="MMVAMWSFLYGLGVAAAFALATPADMRSOSIVELLDREFARTD
GSTATCNFADKRCGGMOGITDKLDYIOGFTAIWTPYTAOLPOTTAAGDYHG
VMQOQIYSLNEWGTADDKALSSALHREGWYLMVDVVAHMGYDAGASSVDSYVEKP
FSSQDFHFPCIONTERDOTQVEDCWLGNTYSLLDIPTRKYVKNEMYSIVSN
YSIDGRIDTVKHOVDKFWPGINKAAGVTCIGEVLDGDPATPCPTQNMQGVLANPIT
YPLNLFKSTGSMDDLNMINTVKSDCDSTLLGTFVENHDNPRFASYTNDIALAKN
VAAFILINDGIPILYAGQSOHYAGNDPANDREATVLSGPTDSELYKLASANAIRNY
AISKDFEYTYKNMPTIKDDTTIARKKSGDSQIYVILSNKASGDSYTLISGAGYT
AGQQLFEVIGCTTVMYSGDGNPVPAGGLPRLVLPTEKTLASICSSS"
join(677..781,837..875,961..1076,1146..1254,1323..1551,
1610..1772,1838..1984,2050..2290,2370..2654)
/gene="amy3"
/product="alpha-amyase"
mat_peptide

intron 782..836
/gene="amy3"
/number=1
837..875
/gene="amy3"
/number=2
876..960
/gene="amy3"
/number=2
961..1076
/gene="amy3"
/number=3
1077..1145
/gene="amy3"
/number=3
1146..1254
/gene="amy3"
/number=4
1255..1322
/gene="amy3"
/number=4
1323..1551
/gene="amy3"
/number=5
1552..1609
/gene="amy3"
/number=5
1610..1772
/gene="amy3"
/number=6
1773..1837
/gene="amy3"
/number=6
1838..1984
/gene="amy3"
/number=7
1985..2049
/gene="amy3"
/number=7
2050..2290
/gene="amy3"
/number=8
2291..2369
/gene="amy3"
/number=8
2370..>2657
/gene="amy3"
/number=9

BASE COUNT 866 a 783 c 715 g 782 t
ORIGIN

Query Match 57.2% Score 992.4; DB 8; Length 3146;
Best Local Similarity 74.8% Pred. No. 3.4e-223;
Matches 1617; Conservative 0; Mismatches 1; Indels 544; Gaps 8;

QY 1 tcaactcaagctctccctctctgacacataaacccacagaagcattttagtggc 60
|||||
DB 564 TCACATCAAGCTCTCCCTCTCTGACATTAACCCACAAAGCAATTATGATGTCG 623
QY 61 cgtgtgtctctatcttctgtaagccttcaggtcaggtcagcctgcttgctcaagc 120
|||||
DB 624 CGTGTGCTCTATTCTGTACGCGCTTCAAGTCCGCGCACTGCTTGGCTGCACGC 683
QY 121 ctgcgactgagatcggaatccatttcttcctctcagagatcgatttgaagagcg 180
|||||
DB 684 CTGCGACTGGGATCGGAATCCATTATTTCTCTCACGATCGATTTGCCAAGACGG 743
QY 181 atgggtcagagcttgacttgtaatactgcgagtc----- 216
|||||
DB 744 ATGGGTCAAGCACTCGCACTGTATACTGCGGATCGGGTGTGTTACTACTAGCTT 803
QY 217 -----agaatactgtgtgtgaactgtgcagggc 245

DB	804	TCAGAAAGAGATGTAACACTGATATAGAAATACCTGATGATGGAACATGGCAGGC	863
QY	246	atcatcgaca-----	255
DB	864	ATCATGACAGCAAGTAATTGGCCCTTTATCAAAAAAAGAAAGAAACGACAGAAAA	923
QY	256	-----agttgagctatccagggaatgg	280
DB	924	TAAATATAAAGAACTCTAGTCCTAACCATCATATAGTTGGACTATATCCAGGGAATGG	983
QY	281	cttcaagacccatctggatcaccocgttaagaccagctgcgcccagaccacgcatalgg	340
DB	984	CTTCACAGCCATCTGATACCCCGCTTACAGCCAGCTGCCACAGACACCGCATATGG	1043
QY	341	agatgcctaccatgctgactggcaagatat-----	373
DB	1044	AGATGCTTACCATGGCTACTGGCAGCAGATATGTAAGTCATTTCTTAATATCTACC	1103
QY	374	-----atactctcgaacgaaa	391
DB	1104	TGTCATCTTTTACATCAATATGACCTAACTTGATGTTTATGATCTCTGTAAGAAAA	1163
QY	392	ctacggcaatgcatgatatgaatgaagcgctctcttcggccctcgaaggggagatgta	451
DB	1164	CTACGGCACTGAGATGACTTGAAGCGCGCTCTTCGGCCCTTCAAGAGGGGATGTA	1223
QY	452	tctatgctcgaatgctgctgaacatat-----	481
DB	1224	TCTTATGGTCGATGTGTGCTTAACCATATGTTGCTGCTTTCGAACTGACTTGGC	1283
QY	482	-----gggctatgaltgagcggtgac	503
DB	1284	GATATGTTTCATTCAGTACGACATGATGATATATCATGAGGCTATGATGAGCGGGTAC	1343
QY	504	tcaatcgattacaagtgtttaaaccgttcagttcccaagactctccaccggttcgt	563
DB	1344	TCAGTCGATACAGTGTGTTTAAACCGTTCCAAAGACTCTTCCACCGCTTCTGT	1403
QY	564	tctatcctaaactatgaagatcgaactcaagltgagaatgctgctgctagagatacaact	623
DB	1404	CTCATTTAAAACTATGAGATCAGACTCAGGTGAGGATCTGCTGCTGAGATTAACCT	1463
QY	624	gtctcttgccctgactcgtatcaccacaagagtgtgtcaagaatgaatggtacgactgg	683
DB	1464	GTCCTCTTGGCTGATCTGATACCAACAAAGATGTGTCAAGATGATGATGATGACACG	1523
QY	684	gtggatcatgtgatacgaactactca-----	711
DB	1524	GTCGATCATTTGTGATTCGAACCTACTCCAGTAAATATTTCTCCCTCATTTCAAACTTGG	1583
QY	712	-----ttgacggtcccgtaicgcacacagtaaaaaact	745
DB	1584	CTGATCGATGATTAATTACGAATATGATGACGGCCTCCGTATTCACACAGTAATAACGCT	1643
QY	746	caagaaggaactctggcccggttacaacaaagccgacagcggttactgcatcggcgaagt	805
DB	1644	CCAGAGGACTTGTGGCCCGGTACAAACAAAGCCGACAGCGCTGTATCTGATCCGCGAGGT	1703
QY	806	gctcgaagtgatccgagcctacactgtccctaccagaagctcagtgagcgctactgaa	865
DB	1704	GCTGACGGTATCCGGCTTACCTTGTCTTACCAGAGCTCTATGAGCGGCTACTGAA	1763
QY	866	ctatcccat-----	874
DB	1764	CTATCCCATGTATGTTCTTCCAAACATGAGCTTCTTTCGAAGTCTCATCTCTTAAGAA	1823
QY	875	-----ttctatccactctcctaagccttcaagttaacctccgacgacatg	920
DB	1824	ACGGCTAAAAACCATGTTATCTATTCACCTCTCAACCCCTTCAAGTCAACCTCCGGAGCATG	1883
QY	921	gaagacctctacaacgatcaacacgcttaaatcgcagctgtccagacttaaacactctg	980
DB	1884	GACGACCTCTTACACATGATATCAACCCGTCAAAATCCGACTGTCCAGACTCAACACTCTG	1943
QY	981	ggcaatctcgtcgaagaaccacgaacccacggttcgttc-----	1021
DB	1944	GGCATTCTCGTGAACACACGACACACCCAGGTTCCTTGTGTAAGTCTTCCCTTTAT	2003
QY	1022	-----ttacaaccaagaca	1035
DB	2004	TTCCGTTCCCATTTTCCACACAGAACCCCACTTAACAAAGACAAAGTTTACACCAAGACA	2063
QY	1036	tagcctcgcgaagaagcgtcgcagatctcatcactcaacgagcgaatcccatct	1095
DB	2064	TAGCCCTCGCAAGAGCTGAGCATTTATTCATCTCTCAACGAGAGATCCCATATCT	2123
QY	1096	acgcgcgcgaagaagcagcactaogcggcggaacgagcccggaacgcgcgaacact	1155
DB	2124	ACGCCGCCCAAGAACACACTACGCCGCCGAAGAACGCCCGCAACCCGGAAGCAACCT	2183
QY	1156	ggctcgggctaccgcagcagcagcgtgtaaaatgtaattgctctcggaacgaca	1215
DB	2184	GGCTCTGGGCTACCCGACCGACAGCGAGCTGTACAAATTAATGCTCCGCGAAGCA	2243
QY	1216	tccggaactatgcccattagaacagatacagatctcgtgactaca-----	1260
DB	2244	TCCGAACTATGCCATTAGCAAAAGATACAGATTCTGTGACTTCAAGATTAAGCAACCT	2303
QY	1261	-----	1260
DB	2304	CTAAGCATCCCTAATGCGCTATCTTACAGATCTGACACAAGAGACTAATCACTGGCA	2363
QY	1261	-----agaacgtgccatctacaagaagacacagatcgccatgycgaagggacagat	1316
DB	2364	ATACAGAACTGGCCCATCTTCAAAAGACACACACATGCGCTGCGCAAGGCGACAGAT	2423
QY	1317	gggtcgaagatcgtgactatcttgcacaagaaggtgtcctcggtgattcgtataccctc	1376
DB	2424	GGGTCGAGATCGTGAATCTTGTCCAAACAAAGGTCCTCGGGTATTCGTAATACCTTC	2483
QY	1377	tccttgatgtgtcgggttacacagcgcgcgaacttgaaggttaattggttcgaag	1436
DB	2484	TCCTTGATGTGTCGGGTTTACACAGCGCGCACAAATTAACGAGGTCAATGGCTGCACG	2543
QY	1437	accgtgacggttgggttcgaatgaatgtgcctgttcctatgycgaagtgagctactaag	1496
DB	2544	ACCGTAGCGGTGTGTCGATGGAATGTGCCGTTCATATGGACAGTGGCTACTCTAG	2603
QY	1497	gtattgatlccgactgagaagatgtgcaggtagaagaatctgtatgactcgtgaaggtg	1556
DB	2604	GTAATGATCCGACTGAGAAAGTTGGCAGGTAGCAAGATCTGTATGCTGTGAAGGGTG	2663
QY	1557	gaagatataatgactgactgacttcaatcctgcatctgacagtgatgattgattgattg	1616
DB	2664	GAGAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2723
QY	1617	ta 1618	
DB	2724	TA 2725	
RESULT 8			
AB021876			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
<p>AB021876 35244 bp DNA linear PLN 17-MAY-2000</p> <p>Aspergillus oryzae amyR, agda, amyA gene for amyR (transcriptional activator), alpha-glucosidase, Taka-amyase A, complete cds.</p> <p>AB021876</p> <p>AB021876.1 GI:7939499</p> <p>amyR, alpha-glucosidase; Taka-amyase A.</p> <p>Aspergillus oryzae (strain:R1B40) DNA.</p> <p>ORGANISM</p> <p>Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.</p> <p>1 (sites)</p> <p>Gomi,K., Akeno,T., Minekoki,T., Ozeki,K., Kumagai,C., Okazaki,N.</p>			

Db	21110	GATATGTTCAATTTCAGTACTGACAAATGAGTAATAATCAAGGCTATGATGAGCGGTAAC		21169
QY	504	tcagtcgattcaagctgcttctaaaccgcttcagttcccaagactactccaccgcttcgt		563
Db	21170	TCAGTCGATTCAGAGTGCTTTAAACCGTTCCAGTTCCCAAGACTACTTCCACCCGTTCTGT		21229
QY	564	ttcattcaaaactatgaagatcagactcagcttgagatctgctgcttgagagataacact		623
Db	21230	CTCATTTCAAAAACCTATGAAAGATCAGACTCAGGTTGAGGATTTCTGGCTGGGAGTAACT		21289
QY	624	gtctccctgctgactctctgatatacccaagatgtgctcaagatgaaatgatacagctg		683
Db	21290	GTCCTCTTGCCGATCTCGATACCAACCAAGAGTGTGTCAAGAAATGATGTGACACTGG		21349
QY	684	gttgagatcatlgtaatcgaactacc-----		711
Db	21350	GTGGGATCATGTGTGATTCGACACTCTCCAGTAAGATATTTCCCATTTCAACACTGG		21409
QY	712	-----ttgaagcgctccgatatcgacacagtaaaacagct		745
Db	21410	CTGATTCGATGATACCTTACGAATTCAGTTAGCGGCTCCGTTATCGACACAGTAAACACAGT		21469
QY	746	ccagaagaagactctggtcccggttaaaaaagccgacagcgctgactctgatatcggaagt		805
Db	21470	CCAGAGAGGACTTCTGGCCCGGTATCAACAAACCCGAGCGGTACTGTATCTGATGCGGAGGT		21529
QY	806	gctcgaagctgataccgagctcaactctgctccatacagaagcgcaltggaacgctgactgaa		865
Db	21530	GCTCGACGCTGATCCGGGCTACACTTTGTTCCTACACAGAACTCATGAGAGGGGTACTGAA		21589
QY	866	ctatcccat-----		874
Db	21590	CTATCCCATGTATGTTGTTCTCCAACCAATGAGCCTTCTTGCAAGTCTCATCTCTTAACGAA		21649
QY	875	-----ttactatccactctctcaacgcttcaagttaaacctccggaagcagt		920
Db	21650	ACGGCTAAAACCAAGTTACATTCCTACCTCCCAACGCGTTCAAGTCAACCTCCGGACGATG		21709
QY	921	gaagcaacttacaacatgatacaaacgctcaaatccagactctgccaagactcaaacctctg		980
Db	21710	GACGACCTTACAACACTGATTCACACACCGCTCAATCCGACTGTCCAGACTCAACACTCTCG		21769
QY	981	ggcacattcgtcgaagaaacacgaacccacaggtctcgctc-----		1021
Db	21770	GGGACATTCGTGGAGAACCAACGACCAACCGAGTTCGTTGTAAGTCTTCCTTTATT		21829
QY	1022	-----ttacaccaacgaaa		1035
Db	21830	TTTCCTGTTCCCAATTTCCACACAGAACCCCACTTAACAMAGAGCAAAATTACACCAAGACA		21889
QY	1036	tagccctcgcgaagaacgctcgaagatctatatactcctaagaacggaatccccatcatct		1095
Db	21890	TACCCCTCGCCAAAGACGTGGACATTCATCTCTCAACGACGGAATCCCATCATCTT		21949
QY	1096	acgcgcgcgaagaacgacatcagccgcggygaaaacgacccgcgaaacgcgcgcgaagcaacct		1155
Db	21950	ACCGCGGCCAAGAACGACACTACGCGCGGGGAAACGACCCCGGAACCGGAAACGAAACCT		22009
QY	1156	gagctcgcggctacccgcacgcagcgaagctgttaagaatlaattgctcctcgcgcgaacgaa		1215
Db	22010	GGCTCTCGGGCTACCCGACCGACGACGACTGTACAAAGTTAAATTTGGCTCCGGCAAGCAA		22065
QY	1216	tccgcgaactatgcatatagcaaaagatacaagatctgttgactaca-----		1260
Db	22070	TCCGGAACTATGCTATTAGCAAAAGATACAGGATTCGTGACTTACAAAGTAAACCAACT		22125
QY	1261	-----		1260
Db	22130	CTTAGCATACCCCTTAATGGCTATCTTCAGAGTATCTGACACAAAGAACTAATCACTGGCA		22189
QY	1261	----agaactgcccactctcaaaagaacgacaaacagatcgccatctgcgaagggcacagat		1316

QY	22190	ATACAGACTGGCCCATCTACAAAGACGACACACAGATGCCATGGCAGAGGCACAGAT	22249
QY	1317	gggtcgcaatcgtagactatctgtccaaagaagtgcttcggtgtagtccgtataccctc	1376
Db	22250	gggtcgcaatcgtagactatctgttcacaaagggtgcttcggtgtagtccgtataccctc	22309
QY	1377	tccttgagtggtgctggtgttaacacagccggtccaggaattgaaggaggttcattggtc	1436
Db	22310	tccttgagtggtgctggtgttaacacagccggtccaggaattgaaggaggttcattggtc	22369
QY	1437	accgtacaggttggttcggaatgtgcctgttcctatagtcagaagtggtggtaccctaa	1496
Db	22370	accgtacaggttggttcggaatgtgcctgttcctatagtcagaagtggtggtaccctaa	22429
QY	1497	gtagtgatccgactgaagaagttggcaggtagaagaatctgtatgctcgtgtaaggtg	1556
Db	22430	gtagtgatccgactgaagaagttggcaggtagaagaatctgtatgctcgtgtaaggtg	22489
QY	1557	gagagctatagtatggtgacgtctatcaatctggtcattgtgaagtaggttgaagtgat	1616
Db	22490	gagagctatagtatggtgacgtctatcaatctggtcattgtgaagtaggttgaagtgat	22549
QY	1617	ta 1618	
Db	22550	TA 22551	
RESULT	9		
LOCUS	ASNTAAG1	2935 bp	DNA linear PLN 27-APR-1993
DEFINITION	A.oryzae Taka-amy1ase A (Taa-G1) gene, complete cds.		
ACCESSION	M33218		
VERSION	M33218.1		
KEYWORDS	Taka-amy1ase A.		
SOURCE	A.oryzae (strain JCM02239) DNA.		
ORGANISM	Aspergillus oryzae		
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Euryotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.		
AUTHORS	1 (bases 1 to 2935)		
	Tsukagoshi, N., Furukawa, M., Nagaba, H., Kirita, N., Tsuboi, A. and Uda, S.		
TITLE	Isolation of a cDNA encoding Aspergillus oryzae Taka-amy1ase A: Evidence for multiple related genes		
JOURNAL	Gene 84, 319-327 (1989)		
MEDLINE	90128276		
FEATURES			
source	Location/Qualifiers		
	1..2935		
	/organism="Aspergillus oryzae"		
	/db_xref="taxon:5062"		
	/db_xref="GI:166531"		
	/note="Taa-G1 mRNA and introns"		
	607..669		
	/note="Taka-amy1ase A signal peptide"		
	<607..772		
	/note="Taka-amy1ase A (Taa-G1) precursor"		
	/number=1		
	join(607..772,828..868,955..1070,1140..1248,1317..1545,1603..1765,1830..1976,2041..2281,2360..2647)		
	/note="Taka-amy1ase A (Taa-G1) precursor"		
	/codon_start=1		
	/protein_id="AAA32708.1"		
	/db_xref="GI:166531"		
	/translation="MMVAVMSLFLYGLQVAAAPALAAPADMSQSIFLLTDRPARDGSTATCNFADOKYCGGTGOGIDIKLDYIOGKGFATWIPVPAOLPOTRAYDADYHGWMODIYLSINENTGIDADDLKALSSALHERHMYFVAVNHMGIDGAGSSVDYSEVPFSSQDFPHFCLIONEDTOYVEDCWIGNTVSLPDLDTKDYVKNRWYMGVSLVNSYSDIARLIDIVKAVQKDFWFGYKRAAGYCTIGEYLDGDPAYTCPCYQWMCVNPYVPLNAFKSTSGSMHDLNMIINVKSCDPSPTLLGTVEVHNDPREASTINDIALAKNVAAFILNDIPIPIYVGOEYHAGVNDPARREKTMASGYPDSLEKYLIASANAIRNYAISKDGEVYIKKMPYIKDDTTIAMRGDGTGSOIVITLSNKGASGDSYTLSSGAGYTAGQQLREVIGCTVTYVGSQGNVVPNAGGLPRVLYPEKLAGSKICSS"		
	join(670..772,828..829)		
	/product="Taka-amy1ase A"		
	mat_peptide		

mat_peptide	join(772,828,.868)
intron	/product="Taka-amylase A"
exon	773..827 /note="Taa-G1 intron A (no splice consensus) ; putative"
intron	828..868 /number=2 /note="Taka-amylase A precursor"
mat_peptide	869..954 /note="Taa-G1 intron B"
exon	join(955,.1070,1140) /product="Taka-amylase A"
mat_peptide	955..1070 /number=3 /note="Taka-amylase A precursor"
intron	join(1069,.1070,1140,.1248) /product="Taka-amylase A"
exon	1071..1139 /note="Taa-G1 intron C"
intron	1140..1248 /number=4 /note="Taka-amylase A precursor"
mat_peptide	1249..1316 /note="Taa-G1 intron D"
exon	join(1317,.1545,1603,.1604) /product="Taka-amylase A"
intron	1317..1545 /number=5 /note="Taka-amylase A precursor"
mat_peptide	join(1545,1603,.1765,1830) /product="Taka-amylase A"
intron	1546..1602 /note="Taa-G1 intron E"
exon	1603..1765 /number=6 /note="Taka-amylase A precursor"
mat_peptide	join(1764,.1765,1830,.1976,2041) /product="Taka-amylase A"
intron	1766..1829 /note="Taa-G1 intron F"
exon	1830..1976 /number=7 /note="Taka-amylase A precursor"
mat_peptide	join(1975,.1976,2041,.2281) /product="Taka-amylase A"
intron	1977..2040 /note="Taa-G1 intron G"
exon	2041..2281 /number=8 /note="Taka-amylase A precursor"
intron	2282..2359 /note="Taa-G1 intron H"
exon	2360..>2647 /number=9 /note="Taka-amylase A precursor"
mat_peptide	2360..2644 /product="Taka-amylase A"
polyA_signal	2784..2789
BASE COUNT	818 a 752 c 657 g 708 t
ORIGIN	

Query Match	57.0%	Score 989;	DB 8;	Length 2935;
Best Local Similarity	74.7%	Pred. M.2.le-222;		
Matches 1613;	Conservative	0;	Mismatches 5;	Indels 541; Gaps 8;
QY	1	tcacatcaagctctccctcctctgacacataaaccacagaagcatltaatgatgcg	60	
Db	557	TCACATCAAGCTCTCCCTCTCTGTGAACATPAACCCACAGAGGCATTATGATGTCG	616	
QY	61	cgtagtgctctctattttgtgaagcgcttgcgagcaactgcttggctgaagc	120	
Db	617	CGTGTGCTCTCTATTCTGTGTAAGCGCTTCAGTCAAGTGCACACTCTTGGCTGCAAGCC	676	

QY	121	ctgggaactgycgatcgcaatccatccatttatcttcctctcaagatcgatcttgcaagacg	180
Db	677	ctgggaactgycgatcgcaatccatccatttatcttcctctcaagatcgatcttgcaagacg	736
QY	181	atggatcgacgactgacgacttgtaactctgacac	216
Db	737	atggatcgacgactgacgacttgtaactctgacac	796
QY	217	-----agaatacttggtgagacatggcaagg	245
Db	797	tcagaaagagcaatataactgacttgatataagaaatctctggtgagacatggcaagg	856
QY	246	atcaatcgaca-----	255
Db	857	atcaatcgacaatgataatgccccctttatcaaaaaaaagaaagaaagcaagaaaaa	916
QY	256	-----atggactatataccaaggaatgg	279
Db	917	atataatataaaaaaactgactgcttaaccatcacatatttggactatataccaaggaatgg	976
QY	280	gcttcaacagccatctggaatcaacccccgcttaaaccccaagctgcccccaagcaacgcatatg	339
Db	977	gcttcaacagccatctggaatcaacccccgcttacaagcccaactgcccccaagcaacgcatatg	1036
QY	340	gagatgctcaccatgagctacttggcaagcaagat-----	373
Db	1037	gagatgctcaccatgagctacttggcaagcaagatataagctgacttcttataatattcac	1096
QY	374	-----atactcttgaacgaaa	390
Db	1097	ctgctcatctttttacatcaatataatgaactaacttgattgatttgaatgactcttgaaacgaaa	1156
QY	391	actaagcgaactgcaagatgaacttgaagagcgtctcttggcccttcataagaggggaatgt	450
Db	1157	actaagcgaactgcaagatgaacttgaagagcgtctcttggcccttcataagaggggaatgt	1216
QY	451	atcttatgctgagatggttgctgaacatat-----	481
Db	1217	atcttatgctgagatggttgctgaacatatatggttctggtcttggcaactgacttgcg	1276
QY	482	-----gggcatagtatgaagggttag	502
Db	1277	ggatattggttcatatttcagtaactgacaaatgagatataatcagaggcctatgataagaggggttag	1336
QY	503	ctcagctgcatcaagtggtgtttaaaccgtttcaagctcccaagactcaactccacccgttctg	562
Db	1337	ctcagctgcatcaagtggtgtttaaaccgtttcaagctcccaagactcaactccacccgttctg	1396
QY	563	tttcaatcaaaaactatgaagatcagactcaggttgaagattgcttggctagaagataaac	622
Db	1397	tttcaatcaaaaactatgaagatcagactcaggttgaagattgcttggctagaagataaac	1456
QY	623	tgctcccttgctgactcctcgataccacaagaagatggttcaagaatgaatgtatagctg	682
Db	1457	tgctcccttgctgactcctcgataccacaagaagatggttcaagaatgaatgtatagctg	1516
QY	683	gggtgactcaattggtatcgaaactactca-----	711
Db	1517	gggtgactcaattggtatcgaaactactcaatccagtaagaatattttctccctcatcttcaacaacttg	1576
QY	712	-----ttgaagcgtcccgatatacgacaacagataaaacagt	745
Db	1577	gctgatcatgatattttwagaaatcagttgagcggcctcggtatwgcacacagtaaaaaacagt	1636
QY	746	ccagaagaagactcttgccccgggtataaanaagccgacaggtgttactgtatctcgycagagt	805
Db	1637	ccagaagaagactcttgccccgggtataaanaagccgacaggtgttactgtatctcgycagagt	1696
QY	806	gctcagaaggtatatacgagctcaacttgcctatacagaagctatgagacggcgtatcagaa	865
Db	1697	gctcagaaggtatatacgagctcaacttgcctatacagaagctatgagacggcgtatcagaa	1756
QY	866	ctatcccat-----	874

QY	112	CTGAAGGCTCCGATCGACAGTAAACAAGT	745
Db	1577	GCTGATCGATGATCTTTACGAAATCAGTTGACGGCTCCGCTATCGACACAGTAAACACGCT	1636
QY	746	CCAGAAGGACTCTGTCGCCGGTACAAACAAGCCGACGGTGTACTCTATCGCGAGGT	805
Db	1637	CCAGAAGGACTCTGTCGCCCGGGTAAACAAGACCGCAGCGCTGTACTCTATCGCGGAGGT	1696
QY	806	GTCGACGGTGTATCCGCGCTACTACTTGTCCATCCAGAAGTCATGTGAAGGCTACTGTAA	865
Db	1697	GTCGACGGTGTATCCGCGGCTCACTTGTCCCTTACCCAGAAGCTCATGTGACGGCGTACTGAA	1756
QY	866	CTATCCCAT-----	874

```

|||||
Db 1757 CTATCCATGTATGTTCTCCAAACATGAGCCCTTCTTGCAAGTCATCTCCTAACGAA 1816
OY 875 -----ttactatocactctctcaagccttcaagcttcaagcttcggcagatcg 921
Db 1817 AGCGTAAACACAGTTACTATCTACTCTCTCAACGCTTCAAGTCAACTCCGGCAGCATGC 1876
OY 922 aagacacttacaatgatatacaacacgltcaaatccgactgttccagactcaacactctcg 981
Db 1877 ACGACCTCTACACATGATGATCAACCGCTCAAAATCCGACTGTCCGACTCAACACTCTCG 1936
OY 982 gcaacatcgttcgagaacacagacaaaccccaeggttcgttc----- 1021
Db 1937 GCACATTCGTCTGAGAACACACAGCAACCCACGGTTCGGTCTGTAGTCTCTCTTATTTT 1996
OY 1022 -----ttacccaacagacata 1037
Db 1997 TCGTTCCCAATTTCCACACAGAACCCCACTAACAAAGCAAAAGTTACACCAACGACATA 2056
OY 1038 gccctcgcaagaacgltcgacgactatcatcctcctcaacgagacatcccatcatctac 1097
Db 2057 GCCCTCGCAAGAAAGTGTGCGAGATTCATCTCTCAACGACGGAATCCCATCATCTAC 2116
OY 1098 gccgagcaagaacgacgactacgcccggcggaacagaccccgcgcaacgcggaagcaactgg 1157
Db 2117 GCCGGCCAAAGAACAGCACTACGCCGGGGAAGACCCCGGAACCGCGAAGCAACCTGG 2176
OY 1158 ctctcgagctcccgacgacgagcgagctgtfacaaagttaattgctctcgcggaagcaatc 1217
Db 2177 GCTTCGGGCTACCCGACCGAGACGAGCTGTACAAAGTTAATTTGCTCGGCAACGCAATC 2236
OY 1218 cgaacacttgcattagcaagaatcagatctgtgactaca----- 1260
Db 2237 CGGAACATTCGCATTAGCAAAAGATACAGAGATTCTGAGACTTCAAGGTACGACAAACCTCT 2296
OY 1261 ----- 1280
Db 2297 AAGCATACCTTAATGGCCTATCTTCAGAGTATCTGACACAAAGACTAATCACTGCAATA 2356
OY 1261 -aagaacttggccatctataaagaacgacacgactcgcacgccaaggagacagatcg 1319
Db 2357 CAGAACTGGCCCATCTTACAAAGACACACAAAGATCGCATGCCAAGGCAAGGCAAGATGG 2416
OY 1320 tgcagactcgtgactatctgttccaacaagggtgttcgggtgacttgcgtatccctcc 1379
Db 2417 TCGCAGATCGAGACTATCTTGTCCAAACAAGGAGTGGTCTGGGTGATTCGTATACCTCTCC 2476
OY 1380 ttgagtggtcggtgttacacagcgcgccagcaaatgtgacgaggtgactgtgacgac 1439
Db 2477 TTGAGCTGGTGGGGTTTACACAGCCGCGCAGCAATTTGACGGAGGTGCTGACAGCAC 2536
OY 1440 gtgagaggtgtgttgcagtggaatgtgcgtttccctatgagcgaggttgagctaacggt 1499
Db 2537 GTGACGGTGTGTTCCGATGGAATGTGCTGTTCTTATGCGAGGTGGCTACTTAAGGTA 2536
OY 1500 ttgtatccgactgagaagttgcaaggttagcaagatctgttagtctgtgaagggttgag 1559
Db 2597 TTGTATTCGAGTGAAGAGTGTGCGAGGTAGCAAGATCTGTAGAGCTGTGAAGGAGTGG 2656
OY 1560 agtataatggttactgtctatcaatctggaattggaagcgaggtttgagtttatgta 1618
Db 2657 AGTATATGATGTACTGCTATTTCATCTGCGATTGGACAGTGAAGTTTGAAGTTATGTA 2715

```

```

RESULT 10
AF139925 AF139925 2970 bp DNA linear PLN 01-DEC-1999
DEFINITION Aspergillus flavus alpha-amylase (amy1) gene, complete cds.
ACCESSION AF139925
VERSION AF139925.1 GI:6492298
KEYWORDS
SOURCE Aspergillus flavus.
ORGANISM Aspergillus flavus

```

```

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE
AUTHORS
Fakhoury,A.M. and Woloshuk,C.P.
TITLE
Amy1, the alpha-amylase gene of Aspergillus flavus: Involvement in
afatoxin biosynthesis in maize kernels
JOURNAL
Phytopathology 89, 908-914 (1999)
REFERENCE
AUTHORS
Fakhoury,A.M. and Woloshuk,C.P.
TITLE
Direct Submission
JOURNAL
Submitted (01-APR-1999) Botany and Plant Pathology, Purdue
University, 1155 Lilly Hall, West Lafayette, IN 47907, USA
FEATURES
source
1. 2970
/organism="Aspergillus flavus"
/strain="86-10D"
/db_xref="taxon:5059"
join(<433..600,656..694,785..900,970..1078,1147..1375,
1434..1596,1662..1808,1874..2114,2194..>2481)
/gene="amy1"
/product="alpha-amylase"
<433..>2481
/gene="amy1"
join(433..600,656..694,785..900,970..1078,1147..1375,
1434..1596,1662..1808,1874..2114,2194..2481)
/gene="amy1"
/EC_number="3.2.1.1"
/codon_start=1
/product="alpha-amylase"
/protein_id="AAFI4264.1"
/db_xref="GI:6492298"
/translation="MVAWMSLFLYGLQVAPALAAFPADMRSGSISYFLTLDRARFD
GSTATCNTADRYKCGTMOGIDKLADYIQMGFTALWITPTVROLQPTAYGAYHG
WYOOOYSLNENYGTADDLKALSLAHERGVYLMVDVYAHMGDAGSSVDVSEFEP
FSSODYFHFPLICIONYEDOTQEDWCLDNTVSLPDLDTKDYKNNYDVGSLVSN
YSIDLRIDYVYKHOKDWPBGKNAAGVYCGEYVLDGDPATCYRONMGVLNVPYI
VPLNAFKSTSSMDLYNMINTYKSDCPDSTILCTPENNDRPASYNDILAKN
VAAFIILNDGIPITYAGEOHRTAGSANDANDEATWLSGTPIDSEIKLISANIRY
AISKDTGEVYKKNMPDIYKDDFTIARMKCTDSQIYITLSNAGASDSTLSLGAVT
AGQQLTEVIGCTTVYVSGDNVPPVMAGLPRVLYPPEKLAGSKICSSS"
BASE COUNT 822 a 745 c 669 g 734 t
ORIGIN
Query Match 56.9%; Score 985.8; DB 8; Length 2970;
Best Local Similarity 74.6%; Pred. No. 1.2e-221;
Matches 1616; Conservative 0; Mismatches 2; Indels 549; Gaps 8;
OY 1 tcaatcaagctctccctctctgaaacaataaaccacagaagatttatgagtgc 60
Db 383 TCACATCAACCTCTCCCTCTTGAAACAATAAACCCACAGAAAGGATTATGATGCTG 442
OY 61 cgtgtgagctctatcttcgtacgacctcaaggttcgagcactgcttgcgtcaacgc 120
Db 443 CGTGTGCTCTCTATTTCGTACGGCTTCAGGCTTCGCGCACCTGCTTGGCTGCAACGC 502
OY 121 ctgcgagcttgcagtcgcaatccattatctcccttcaagatcgatttgaaggagc 180
Db 503 CTGGGAGCTGGCATGCAATTCATTATTCCTCTCAAGATGATTTGCAAGAGCAGC 562
OY 181 atggtgtcgaagactgctgactgttaatactgcgagatc----- 216
Db 563 ATGGGTGAGAGCTGGACTTGTAAATACGCGGATGGGCTGTGTTGCTCTACTACTT 622
OY 217 -----agaatactgtgtgtgaacatgacagagc 245
Db 623 TCAGAAAGAGGAATGTAAACTGACTTGATATACAAATFACTGTGCTGGAACATGGCAGGCG 682
OY 246 atcatcgaca----- 255
Db 683 ATCATGCAAGTAAGTAATTCCTCCCTTATCAAAAAAAGAGAGAGAGAGAGAGAG 742
OY 256 -----agttgactatataccagga 275

```

```

Db 743 AAAATAAATAAAGAAAGAACTAGTCTAACCATCATATGAGTATATCCAGGA 802
Qy 276 atgggttcaacagcatcttgatcacaccccggttacagcccaagctgcacacacgca 335
Db 803 ATGGGCTTCACAGCATCTGGATCACCCTGTTACAGCCAGCTGCCCCAGACACGCA 862
Qy 336 tatggagatgacctacatggtctactgcaagcaagatat----- 373
Db 863 TATGAGATGCTCAACCTAGGCTACTGGCAGCAGATATGTAAATGATTTCTTAAATAT 922
Qy 374 -----atacctcttgaaac 386
Db 923 CTACCTCTCATCTTTTACATCAATMGAACTAACTTGATGGTTTAAAGATCTCTCGAAC 982
Qy 387 gaaatacagcagcagatgagaatgacttgaagcgctctcttcgcccctcatgagaaggg 446
Db 983 GAAAACTAGGCACTGAGATGACTTGAAGGCGCTCTCTCGGCCCTTCATGAGAGGGG 1042
Qy 447 atgfatcttatggtcgatggtgtgttaaacatat----- 481
Db 1043 ATGATATCTTATGGTCAGTGTGTGCTAACCATATATGTTCTGTGCTTTCGAACTGACT 1102
Qy 482 -----gggctatgatgagcg 498
Db 1103 TCGCGATATGATTGATTTCAGTACTGACAAATGATATATCGGGCTATGATGAGCGG 1162
Qy 499 gtatgctcaatcgatatacagtgtgtttaaacgcttcacgttcccaagactctccaccgt 558
Db 1163 GTACTCTAGTCGATTAAGTGTGTAAACCGTTCAGTTCGCCAAGACTCTCTCCACCCCT 1222
Qy 559 tctgttccatccaacatagaagatcagatcaggttgaagatgctggtcagagagata 618
Db 1223 TCTGTCTCATTTCAAACTATGAAGATCAGACTAGGTTGAGATTCCTGTGCTAAGAGATA 1282
Qy 619 acactgtccctcgctgctgatactcgaatacaacaaagatgtgtgcaagaaatgaatgtacg 678
Db 1283 ACACGTGCTCTCTGCTGATCTCGATACCCAAAGATGTGTCAAGATGATGTGTACG 1342
Qy 679 actggtggtgagatcttgatcgaactactca----- 711
Db 1343 ACTGGGTGGATCTTGTGATCGAATCTCCAGTAAGATATTTCCCTCATTTCTACAA 1402
Qy 712 -----ttgacggtcccgatcgaacagataaa 740
Db 1403 CTTGGCTGATGATGATTAAGTAATCAGTTGACGGCTCCGTATCGACACAGTAAAA 1462
Qy 741 caggtcgaagaagacctctgcccgggtacaaagaagccgagcggtgactgatatcgcc 800
Db 1463 CACGTCCAGAGGACTTCTGTGCCCCGGGTACAAAGCCGCAAGCGGTACTGTATCGGC 1522
Qy 801 gaagtgtcgaagtgatccgagcctaaactgtccctacagaagcactgaagcagcgata 860
Db 1523 GAGGTGCTCGACGCTGTGATCCGGCTACACTTGTCCCTACAGAAACGATGAGCGCGTA 1582
Qy 861 ctgaacataccat----- 874
Db 1583 CTGAACATATCCATGTATGTCTCTCCAAACGATGAGCGCTTCTTGCAAGTCTCATCTCCGA 1642
Qy 875 -----ttactatcactctcctcaacgctcctcaagtcaactccggca 915
Db 1643 ACGAAAGCGGTAAACAGATTAATCACTCTCAACAGCGCTTCAAGTCAACTCCGGGA 1702
Qy 916 gcatggaagacacttacaacatgatcaaacacggtcaaatcgaactgtccagactcaaac 975
Db 1703 GCATGAGAGACCTCTACAAATGATTAACACCGTCAAAATCCGACTGTCGAGTCAACAC 1762
Qy 976 tccgtgggacactgtcgaagaacacgaacacacggttcgcttc----- 1021
Db 1763 TCCTGGGCACTTCTGTGAGAACCAACCAACCCAGGTTCTGTGTAAGTCTTCCCTT 1822
Qy 1022 -----ttacaccaa 1030

```

```

Db 1823 TTATTTTCCTTCCCAATTTCCACAGAAACCCCATTAACAAGAGCAAGTTACACCA 1882
Qy 1031 cgaatagccctcgcaagaagctgcgaatcatcatcctccaaagacgaatcccat 1090
Db 1883 CGAATAGCCCTCGCCAAAGACCTCCGACATTCATCATCTCAAGACGAGATCCCAT 1942
Qy 1091 catcagccgccaagaacagactacgcgcgcggaagaaccccgaaacgcgaagc 1150
Db 1943 CATCTACGCGCGCAAGAAACAGCATACGCGCGGAAAGCAACCCGCAACCGCGAAGC 2002
Qy 1151 aacttgctctcggtctaccacgcagacgcagcgtgatacaagttaattgctccgaa 1210
Db 2003 AACCTGCTCTCGGGCTACCCGACGACGACGAGCTGTCAAGTTATTTGCTCCGCA 2062
Qy 1211 cgcacccggaactatgcatatgaagaatagcagatcgttgacctaca----- 1260
Db 2063 CGCAATCGGAACTATGCCATTAGCAAAAGTACGAGATTCTGACCTACAAAGTTAAGC 2122
Qy 1261 ----- 1260
Db 2123 AACCTTAGCATACCTTAATGGCCTATCTTCAGACTATCTGACACACAGACTAATCAC 2182
Qy 1261 -----agactggcccatctacaagaacgacacacagatcgcaatgcgaagga 1311
Db 2183 TGCAATACAGAACTGCGCCATCTACAAGACGACACAAAGATCGCATCGCAAGGCA 2242
Qy 1312 cagatggtcgagatcgtgactatcttgcacaagaaggtgcttgggttgatgata 1371
Db 2243 CAGATGGGTGCGAGATGTACTATCTTGCAACAAGGTCCTTGCGGTGATGCTATA 2302
Qy 1372 cccctccttgatggtgtgcggttacacagccgcgcaagaatctgaacgagatcgt 1431
Db 2303 CCTCTCTTGAAGTGTGCGGGTTACACACCGCGCAAGCAATTAAGGAGATTTGGGT 2362
Qy 1432 gcaagccgtgaagtggtgtcgtgatagaatgtgcctgtcctatgcaagatggtgctac 1491
Db 2363 GCACGACCGGTGACGTTGTTGCGATGGAATGTGCTTCTTATGAGAGGTGGCTAC 2422
Qy 1492 ctgaagtattgatacgcagcagaaagtgtgcaggtgaagaatctgatactgacgtga 1551
Db 2423 CTAGGTATTTATCCGACACTGAGAAAGTTGCAAGATCTTATAGCTGTGTA 2482
Qy 1552 ggtgtgagatlatatgagtbaetgclatcaactcgtgcattgagacgtgaattgaatt 1611
Db 2483 GGTGGAGAGATATGTAGTACTGATTAATCTGGCACTTGACAGTGTGAGTTGAGTT 2542
Qy 1612 tgatgta 1618
Db 2543 TGATGTA 2549

```

RESULT 11
ANAMBY 2443 bp DNA linear PLN 01-DEC-1992
LOCUS A.niger amyB gene for alpha-amyase.
ACCESSION X52756
VERSION X52756.1 GI:2325
KEYWORDS alpha-amyase.
SOURCE Aspergillus niger
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1 (bases 1 to 2443)
AUTHORS Korman,D.R., Bayliss,F.T., Barnett,C.C., Carmona,C.L., Kodama,K.H., Royer,T.J., Thompson,S.A., Ward,M., Wilson,L.J. and Berka,R.M.
TITLE Cloning, characterization, and expression of two alpha-amyase genes from Aspergillus niger var. awamori
JOURNAL Curr. Genet. 17 (3), 203-212 (1990)
MEDLINE 90254827
FEATURES Location/Qualifiers
source 1..2443 /organism="Aspergillus niger"/strain="var. awamori, UVK143f"

Db 871 GATATGCTTCATTTACGATCTGACAAATGAGTAAATATCAGGGCTATGATGAGCGGGTACG 930
 Qy 504 tcaagtcgatatacagtggttlltaaacggttccagttcccaagactacttccaccggttcgt 563
 Db 931 TCAGTGTGATTTACAGTGTGTTAAACGGTTCAAGTTCACCAAGACTACTTCCACCCTTCGT 990
 Qy 564 ttcatacaaaactatgaagatacagactcaggttgaagattgtcgtgtaagataact 623
 Db 991 TTCATTCAAAACTATGAAATGAGACTAGACTAGGATTTGCTGGCTAGAGATTAACACT 1050
 Qy 624 gtctcttcgcttgatctgatcccaagagatggtgtaaaagaatggttaagactggtg 683
 Db 1051 GTCTCTTGCTGCTGATCTGATTCATACACCAAGAGATGTGCTCAAGATGAATGTACAGCTGG 1110
 Qy 684 gtggagatcatgtgtatcgaactactcca----- 711
 Db 1111 GTGGGATCATTTGGTATCGAATACCTCAGTAAGATTTTCTCCCTCATTTCTACAACTTGG 1170
 Qy 712 -----tgaagcgctccgtatcgaacaagtaaacagct 745
 Db 1171 CTGATCGATGATTAATACGAAATCAGTTGACGGCTCCCTATCGACACAGTAAACAGCT 1230
 Qy 746 ccagaagagactctgagccgggtatacaaaagccgcaagcggtgtactgtatcgcgaggt 805
 Db 1231 CCAGAAGGACTTCTGGCCCGGGTACACAAAGCCGAGCGGTGTACTGTATGCGCAGGT 1290
 Qy 806 gctcgaagtgatccggtctactactgttccctacagaacgctatgagagcggtactgaa 865
 Db 1291 GCTCGAGGTGATCCGGCTTACACTGTGCTCCACAGAAAGTCATGGAAGGGGTACTGAA 1350
 Qy 866 ctatcccat----- 874
 Db 1351 CTATCCCATGTATGTTCTCCAACCCAGGCTTCTTGCAAGTCTCATCTCTTAACGA 1410
 Qy 875 -----ttaatacactcctcaagccttcaagttaactcctcgcgagctg 920
 Db 1411 ACGGCTAAACAGTTACTATCTCACTCTCTCAACGGCTTCAAGCTCAACCTCGGCGAGCTG 1470
 Qy 921 gaagactctacaatatgatacaacggttcaaatcgaactgtcgaactcaaacactcgt 980
 Db 1471 GACGACTCTTACACATGATGATCAACCGTCAAACTCGACTGTCCGAGCTCAACACTCTGT 1530
 Qy 981 ggcacattcgtctgagaacacagacacacacaggttcgtc----- 1021
 Db 1531 GGCACATTTGTTGAGAACCCACAGAACCCACGTTGCTTCTTAAGTCTTCCCTTTATT 1590
 Qy 1022 -----taagcaaacgaca 1035
 Db 1591 TTCGGTTCCTCAATTTCCACACAGAACCCACCTTAACAAAGACCAAACTTACACCAACGACA 1650
 Qy 1036 tagccctcgcaagaacgctcgacatcatcatcctcaacgaagaaatcccatcatct 1095
 Db 1651 TAGCCCTCGCCCAAGAACGTCGACGATTCATCTCTCAAGACGGAATCCCATCATCT 1710
 Qy 1096 agcgcgcccaagaacagactacgctcgcggaagaaagaccccggaacgcggaagcaact 1155
 Db 1711 ACGCGCGCCAAACAGACTAGCGCGGCGGAAAGACCCCGCAACCGCGAAGCAACCT 1770
 Qy 1156 agctctcggtctaacccgagccgagcggtgtacaagtttaattgtcctcgcggaagca 1215
 Db 1771 GGCTCTCGGGCTACCCGACCGACGACGAGCTGTACAAATTATTTGCTCCGGAACGCAA 1830
 Qy 1216 tccggaactatgcatatagcaagaatcgtgtacttaca----- 1260
 Db 1831 TCCGGAACATGCTAGCAAGATACAGATTCGTGACTACAAAGTAAAGCAACACT 1890
 Qy 1261 ----- 1260
 Db 1891 CTAAGCATACCTTAATGGCTATCTTCAAGATATCTGACACAGAGACTTAATCACTGSCA 1950
 Qy 1261 -----agaactggcccatctacaagaacacacacagatcgcgcattcgcaagggcacagat 1316
 Db 1951 ATACGAAGTGGCCCATCTACAAAGACACACACAGATCCGATGCGCAAGGGCACAGAT 2010

Qy 1317 ggttcgacagatcgtgactatctgtccaagaaggtgtcttcggtgatctgtataccctc 1376
 Db 2011 GGGTCGCGAGATCGTACTATCTTGTCCAACAGGGTCTTGCGGTGATTCATACCTTC 2070
 Qy 1377 tcccttgagtggtcgggtttacacagccgccaagcaatttgaacgaggtcatctggtcagc 1436
 Db 2071 TCCTTGAGTGTGGCGGTTACAGAGCCGCCAGCAATTGACGAGGATTCATTGGCTGACG 2130
 Qy 1437 accgtgacggttgggttcggatgaaatgtgcgttccctatggaagtggtgactaagg 1496
 Db 2131 ACCGTGACGCTTGGTGGATGAGAAATGTGCTCTTCTTAATGCGCAGGGGCTACAGG 2190
 Qy 1497 gttatgtaccgactggaaggttgagcggtgaagaaatcgtgtagctcgttaagaggtg 1556
 Db 2191 GTATTGATTCGAGCTGAGAACTTGCGAGTACGCAAGATCTGTACTACTGCTGAAGAGTG 2250
 Qy 1557 gagaatataatgtagtactgctatcaactcgtgcatctgagcagtgagttgagttgatg 1616
 Db 2251 GAGAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2310
 Qy 1617 tac 1619
 Db 2311 GAC 2313

RESULT 12

ANMYA
 LOCUS ANMYA 2520 bp DNA linear PLN 01-DEC-1992
 DEFINITION A.niger amya gene for alpha-amyase.
 ACCESSION X52755
 VERSION X52755.1 GI:2323
 KEYWORDS alpha-amyase.
 SOURCE Aspergillus niger.
 ORGANISM Aspergillus niger.
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

REFERENCE

1 (bases 1 to 2520)
 Korman,D.R., Bayliss,F.T., Barnett,C.C., Cazmona,C.L., Kodama,K.H., Royer,T.J., Thompson,S.A., Ward,M., Wilson,L.J. and Berka,R.N. Cloning, characterization, and expression of two alpha-amyase genes from Aspergillus niger var. awamori
 JOURNAL Curr. Genet. 17 (3), 203-212 (1990)
 MEDLINE 90254827

FEATURES

Source

1..2520
 /organism="Aspergillus niger"
 /strain="var. awamori, UVK143f"
 /db_xref="taxon:5061"

CAAT_signal

TATA_signal

gene

exon

CDS

/number=1
 /join(201..368,424..462,548..663,733..841,910..1138,
 1197..1359,1425..1571,1637..1877,1997..2241)
 /gene="amyA"
 /EC_number="3.2.1.1"
 /function="hydrolysis of alpha-1,4-glucosidic bonds in
 polysaccharides"
 /note="1, 4-alpha-D-glucan 4-glucanohydrolase"
 /codon_start=1
 /product="alpha-amyase-precursor"
 /protein_id="CA3696.1"
 /db_xref="GI:2324"
 /db_xref="SWISS-PROT:Q02905"
 /translational="MMVAMWSLEFLGLOVAAPALATPADMRSGSYFLNDRFARTD
 GSPATACNADOKYGGTMOGILDKLDYLOGGFTATITPTNOLPOTTAAGAYHG
 YMOODYSINENYGRADDKALSSALHEGKNTLMDVYANHMGTDGAGSSVDYSYFK
 FSSDYHFHCFCTONEDQVEDEWLGNTVSLPDLDTTKDQVYENYDWGSLVSN
 YSIDGRLIDTVKHQKDFWPGYNKAGAYCIGSEVLDDGDPATCPYQNVMDGLNPIY

sig_peptide
 mat_peptide
 /gene="amyA"
 join(264..368,424..462,548..663,733..841,910..1138,
 1197..1359,1425..1571,1637..1877,1957..2238)
 /gene="amyA"
 /EC_number="3.2.1.1"
 /note="1,4-alpha-D-glucan 4-glucanohydrolase"
 /function="hydrolysis of alpha-1,4-glucosidic bonds in
 polyglucosans"
 /product="alpha-amylose"
 369..423
 /gene="amyA"
 /number=1
 424..462
 /gene="amyA"
 /number=2
 463..547
 /gene="amyA"
 /number=2
 548..663
 /gene="amyA"
 /number=3
 664..732
 /gene="amyA"
 /number=3
 733..841
 /gene="amyA"
 /number=4
 842..909
 /gene="amyA"
 /number=4
 910..1138
 /gene="amyA"
 /number=5
 1139..1196
 /gene="amyA"
 /number=5
 1197..1359
 /gene="amyA"
 /number=6
 1360..1424
 /gene="amyA"
 /number=6
 1425..1571
 /gene="amyA"
 /number=7
 1572..1636
 /gene="amyA"
 /number=7
 1637..1877
 /gene="amyA"
 /number=8
 1878..1956
 /gene="amyA"
 /number=8
 1957..2241
 /gene="amyA"
 /number=9
 2242..2370
 /number=9
 2370..2375
 BASE COUNT 682 a 645 c 581 g 611 t 1 others
 ORIGIN
 Query Match 52.4%; Score 908.2; DB 8; Length 2520;
 Best Local Similarity 73.6%; Pred. No. 2.3e-203;
 Matches 1537; Conservative 0; Mismatches 8; Indels 544; Gaps 8;
 QY 1 tcacatcaagctctccctctctgaacaataaaccacagagaagcattatgatgtcg 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 151 TCACATCAAGCTCTCCCTCTCTGTAACAATAAACCCACAGAGGCAATTATGATGTCG 210
 QY 61 cctgtgtgtctctatattctgtacaggccttcaagtcgcggcaactgcttggctgcaagc 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 211 CGGTGGTCTCTATTTCTGTACGGCTTCAAGTCGGCCACCTGCTTGGCTGAACGC 270
 QY 121 ctgcgactctgcagatcgcaatccattatcttccctcagagatcgatgtgcaaggacg 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 271 CCGCGAGCTGGGAGATCGCAATCCATTTATTCCTTCACGAGTGTGCAAGACGG 330
 QY 181 atgggtcgacgactgcgactgttaactctgagatc----- 216
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 331 AAGGTCACAGACTGGGACTGTGAATACGGGATCAGGTGTGTTACTACTAGCTT 390
 QY 217 -----agaactctgtgtggaacatgycaggc 245
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 391 TCAGAAAGAGATGTAACTGACTTGATATGAATATCTGTGGTGAACATGGCAGGGC 450
 QY 246 atcatcgaca----- 255
 ||||||||
 Db 451 ATCATCGACAGGTAAATGCCCCCTTATCAAAAAAAGAGAAAAAGCAAGAAAAA 510
 QY 256 -----agttgagctatatccagggaatgg 280
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 511 TAAATTAAGAACTCTAGCTTAACATCACATGATGTGACTATATCCAGGGAATGG 570
 QY 281 ctccacagccatctgtatcaccccgctacagccagctgcccacagccacgcatatgg 340
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 571 CTTCACAGCATCTGTGATCACCCCGCTTACAGCCAGCTGCCACAGCCGCAATAGG 630
 QY 341 agatgcctacatgctctactgycagcagatat----- 373
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 631 AGATGCTACACAGCTGCTACTGGCAGCAGATATGTAGTGAATTCCTTAAATATCTACC 690
 QY 374 -----atactctgaagcaaa 391
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 691 TGTCAATCTTTTACATCAATATGACTAATTGATGCTTTAGATACTCTCTGACAGAAA 750
 QY 392 ctacgacactgcagatgacttgaagcgctctcttcgcacctcatgagagggaatgta 451
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 751 CTACGGGACTGCAGATGACTTGAAGGCGCTCTCTGCGCCCTTCATGAGAGGGGATGTA 810
 QY 452 tcttatgctcagatgctgttgctacatat----- 481
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 811 TCTTATGCTGATGTGTTGCTTAACCATATGTTGTTGCTTGTGCAACTGACTTGGC 870
 QY 482 -----ggctatatgagcgggtagc 503
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 871 GATATGTTTATTCAGTACTGACAAATGAGTAAATATCAGGGCTATATGAGCGGTAGC 930
 QY 504 tcaatcgattacagatgtgtttaacggttcagttcccaagactactccaacggtctgt 563
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 931 TCAGTCGATTACAGTGTGTTAAACCGTTTCAGTCCCAAGACTACTCCACCGTCTGT 990
 QY 564 ttcattcaaaactatgaagtcagactcaggttgaagattgctcgggttagagataact 623
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 991 TTCAATCAAAACTATGAAGATCAGACTCAGGTGAGGATGCTGGCTAGGAGTTAACACT 1050
 QY 624 gtctccttgctgctcgtatccatccaaagagatggtcaagaatgaatgtaacgactg 683
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1051 GTCTCTTGCTGCTGATCTCGATACACCAAGAGATGTGTCAAGATGAATGTACGACTGG 1110
 QY 684 gtggagcattglatcgaaactactca----- 711
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1111 GTGGATCATTTGATGAGACTACTCCAGTAAGATATTTCTCCCTATTTCTACACTGG 1170
 QY 712 -----ttgacgctctccgfatcgacaagtaaaacagt 745
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1171 CTGATCGATGATACTTACGAATCAGTTGACGGCTCTCCGATGACACATTAACACGT 1230
 QY 746 ccagaaggaactcttggccgggttacaacaagacgacgaggttactgttgcgcaaggt 805
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1231 CCAGAGGACTTCTGGCCCGGTACAAACAGCCGACGGCTGTACTGTATGCGCAGGT 1290

Qy 421 tctctcgcccttcacagagagggatgatactatgctcagatggtgttgctaacata 480
Db 371 TCCTCGAATGCTCTTCAGACCCCGGGAATGTAACCTCATGTCGAGCTGTCCCTAACACA 430
Qy 481 tggcctatgatgagcggtgtagctcagtgattacagtggtttaacccgttacc 540
Db 431 TGGGCTACGGCAGGTACGGGCAACGATGTGATACAGCGTCTTCGACCCCTTGACCTCT 490
Qy 541 aagcctcttcacacgcttctgttcttcaaacctatgaagatcagatcagatgtag 600
Db 491 CCTCTACTTCTCATCTCCTACTGCTCTCATACAGATTGGGACACTTACCATGCTCCAG 550
Qy 601 attgctgtagagataaacactgltctcctgctcgaatctcgaatacacaagatgtag 660
Db 551 ACTGTGGGAGGAGTGACACCATGCTGCTCTGCCAGATCTGAACACACGAAACGCCCG 610
Qy 661 tcaagaatgaatggttaagactcgtggttgatcatatggtatgtaacttcaatgtag 720
Db 611 TGAAGAACCATTTGGTAGCATTTGGGTACCCGACCTGATCCAACTACTCAGTCAGCGCC 670
Qy 721 tccglatcagacagtaaaacagctccagaagagactctgcccgggttacaacaacg 780
Db 671 TCCTGATGACAGTGTGGAAGAAATCCAGACCCGACTTCTTCCGGGCTTACCAAGACAG 730
Qy 781 caggcgtgtactgtatcggcgaaggtgctcgaaggtgtagctcggcctaacactgtcc 840
Db 731 CAGGAGTCTACTGCGTGGTGAAGTTCAGACAGGCAACCTGCTCTGACAGCCCATACC 790
Qy 841 agaacgtatggaagcggttactgaactatccattactatccactcctcaagcctca 900
Db 791 AAAAATATCTAGATGTGCTTCTCAACTATCCATCTACTGGCACTCTCTACGCCCTTGG 850
Qy 901 agtcaacctcggcgaagcagctcgaagcctctcaacatgatacaacacgccaatcgact 960
Db 851 AATCTTCAGGGGAGCATTCAGCAACCTCTCAACATGATATAATCCGTCCGACGACT 910
Qy 961 gtccagactcaacactcctcgtggcacaatcgtcgaagcaacacgcaacccaggttgc 1020
Db 911 GCTCGATCCGACCCCTCTGCGCAACTTATATGCAAAACACAGCAACCCCGCTTCCCT 970
Qy 1021 cttaacaaagacatagcctctgcgaagacgtcgcagcatlatacctcgaagc 1080
Db 971 CCFACACATCCGACTACTCCCAACCCAAACGTTCTCAGCTACTCTTCTCCGACG 1030
Qy 1081 gaatcccatcatcttaagccggccaagaacacactaagccggcgaagcaaccccgga 1140
Db 1031 GCATCCCATGTCTACGCCCGCGGCGAAGACAGCACTACTCCGCGCGGACCTGCCCT 1090
Qy 1141 accggaagcaacgtgctcgtcgtcgaacccgcaagcgaagcgtgtagaagtaatg 1200
Db 1091 ACAAGCGAGTACCTGCTATCAGGCTTCAGACACCTCCGCGGAGCTCTACACCTGATG 1150
Qy 1201 cctcgggaagcaactccggaactatgccaatgaagaagatacaggaattgtgaactaca 1260
Db 1151 CCACCAACAAAGCGATCCGGAACCTAGCTATCTCAGCAAGACTCCGACTACATTTACTTACA 1210
Qy 1261 agaatggcccatctacaagaacgacacacacatcgccatcgcgaagcgacagatgtag 1320
Db 1211 AGAAGACACCAATCTACACAGACAGCAACCATTCGCATCGCAAGAGCACTCCGCT 1270
Qy 1321 cgcagatcgtactatctgttccaacaagggtgcttcggtgtagctgtatatacctcct 1380
Db 1271 CCGAATATCATCGCTCTCTCCCAAAAGGCTCTCCGGAAGAGCATACACCTCAACC 1330
Qy 1381 tgaatggtggtgttaacacagccggccagcaattgacggaggtcatatggttgcagcag 1440
Db 1331 TCAGCGGAGGGGCTACAGCTCCGCAAGAGCTCATGAAAGCTTACACCTGCACGCTCG 1390
Qy 1441 tgaaggtgtgtcgtatgtaaatgctcgtgttccatagcgcggtggtcactactgggtat 1500
Db 1391 TGACGCTGAGCTCGAAGCGGAGTATCCTGTGCGGAGTCTTCGGGATTTACCTAGAGTTC 1450
Qy 1501 tgtatcgcactgtagaaggtgtagcaagatcgttag 1540

Db 1451 TCCTCCCTGCTTCGGTGGTGTGATGTTCTTCGCTTGTGG 1490

RESULT 14
LK030376 2239 bp mRNA linear PLN 27-JUN-1997
LOCUS LK030376
DEFINITION Lipomyces kononenkoae subsp. spenceri martinsiae alpha-amy-lase mRNA,
complete cds.
ACCESSION U30376
VERSION U30376.1 GI:1173536
KEYWORDS
SOURCE
ORGANISM
Lipomyces spenceri martinsiae.
Lipomyces spenceri martinsiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Lipomycetaceae; Lipomyces.
REFERENCE
AUTHORS Steyn, A.J., Marmur, J. and Pretorius, I.S.
TITLE Cloning, sequence analysis and expression in yeasts of a cDNA
containing a Lipomyces kononenkoae alpha-amy-lase-encoding gene
JOURNAL Gene 166 (1), 65-71 (1995)
MEDLINE 96105202
REFERENCE 2 (bases 1 to 2239)
AUTHORS Pretorius, I.S.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-1995) Isak S. Pretorius, Microbiology, University
of Stellenbosch, Victoria, Stellenbosch, Western Cape 7600, South
Africa

FEATURES
source
1. 2239
/organism="Lipomyces spenceri martinsiae"
/strain="IGC4052B"
/sub-species="spenceri martinsiae"
/db_xref="taxon:56873"
/lab_host="Saccharomyces cerevisiae"
242..1954
/codon_start=1
/product="alpha-amy-lase"
/protein_id="AAC49622.1"
/db_xref="GI:1173537"
/translation="MCGSTLSASLYVNDYDKITVLTLYLTSSGTTSLALIPVMS
NNELMTLSAIAAGAVETTGASYSDDSTVYTTISLDPLPTTASAPVPTGAAWRR
SLYOVVTRFAFATGDSITVSCDVTDRVYGGSYKGIINMLDYIGKMGTAIVSIVE
NIPDGTGAYVAGHWMKDI FALNFEFGADLI ALAELINRMVLMVLDIVNHF
SGNHADVSEFPYSODYFHFPCIMADYENQTNVEECWLGDSVPLVDVNDLDTV
KSEYOSWKOLIANYSIDGRIDPKYHDMFMVFOEAGIITVGEVDPDPEYTC
YQENLDGLNLPVYVYSAFOROGSSISLVMDITLTKSCIDITTLIGSPLTENDNP
RFPSTSDSLIKNAIATITLSDGILPITITGQEGGLNGNDPYNREALMPGTGISTST
FVEYTLASLNOIRNAIYIDPITLYQNMVITSDSTTLAMRGFTGNQITVLSMGSS
GSSYTLTSLNMGYTAASSVVEILTCTAVTVLDSGLNAVPMGGLPRVFPESQLVSG
ICSM"

BASE COUNT 549 a 488 c 491 g 711 t

ORIGIN

Query Match 25.1%; Score 434.8; DB 8; Length 2239;
Best Local Similarity 56.2%; Pred. No. 8..2e-92;
Matches 817; Conservative 0; Mismatches 637; Indels 0; Gaps 0;

Qy 97 cggcagctcgttctgtcgaacgctcggagctcggatcgcaatccattatccttc 156
Db 504 CCGCGTCTGCTTCAGACAGACAGCTCAAAATGGCAGAGTCTCATCTCATGAGTCCG 563
Qy 157 tcaagatcgtattgcaagaagcagatggttgcagacgtcgcattgatactcggatc 216
Db 564 TAACGTATGATTTGACACGACACTGAGTGCCTCATTAATATTCATGCGAGTACCGGATA 623
Qy 217 agaatcactggttggaacatggaagcagcatcagacagttggaactatccaggaa 276
Db 624 GGGTCTATTTCGAGGGGTCTTACCGGGGATCATCAACTGCTGATTTACATCAAGGA 683
Qy 277 tggcttcaagcagcatctgatacccccgttacaagccagctgccccagacacagcat 336
Db 1331 TGAAGAACCATTTGGTAGCATTTGGGTACCCGACCTGATCCAACTACTCAGTCAGCGCC 670

[illegible]

QY	1417	cggaagtcattggctgcgcacgcacgctgcagcgttgcttgatcgatggaatgtgccttcccta	1476
Db	1824	ATGAGATCTTTCACAGACGACAGCTGTGACCTGTGGATTTTCCGGAAATTTGGCAGAGCCAA	1883
QY	1477	tggcggcgtgggcttacctgaggtatgtatccgacgcaggaagtggcaggtagcaagatct	1536
Db	1884	TGTCGGGTGGCTTACCGAGAGTGTTTACCCCGAATCCGCACTGGTGTGGCTGTGGAATCT	1943
QY	1537	gtagtaagctcgtga	1550
Db	1944	GCTCCATGTATAGGA	1957
RESULT	15		
SCSMA2			
LOCUS		2143 bp	DNA linear
DEFINITION		S.occidentalis SWA2 gene for alpha-amyrase.	PLN 10-AUG-1993
ACCESSION	X73497		
VERSION	X73497.1	GI:396561	
KEYWORDS		alpha-amyrase; SWA2 gene.	
SOURCE		Debaromyces occidentalis.	
ORGANISM		Debaromyces occidentalis	
REFERENCE		Eukaryota, Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Debaromyces.	
AUTHORS		1 (bases 1 to 2143)	
TITLE		Clarios,M.G., Abarca,D., Fernandez-Lobato,M. and Jimenez,A.	
JOURNAL		Molecular structure of the SWA2 gene encoding an AMY1-related	
MEDLINE		alpha-amyrase from Schwanniomycos occidentalis	
REFERENCE		Curr. Genet. 24 (1-2), 75-83 (1993)	
AUTHORS		2 (bases 1 to 2143)	
TITLE		Jimenez,A.	
JOURNAL		Direct Submission	
FEATURES		Submitted (02-AUG-1993) A. Jimenez, Centro de Biologia Molecular,	
source		Universidad Autonoma, Cantoblanco, 28049 Madrid, SPAIN	
		Location/Qualifiers	
		1..2143	
		/organism="Debaromyces occidentalis"	
		/strain="ATCC 26077"	
		/db_xref="taxon:27300"	
		234..1757	
		/gene="SWA2"	
		234..1757	
		/gene="SWA2"	
		/EC_number="3.2.1.1"	
		/codon_start=1	
		/product="alpha-amyrase"	
		/protein_id="CA51912.1"	
		/db_xref="GI:396562"	
		/db_xref="SPRMBL:008806"	
		/translation="MKFATILSTTALASLSLVASKPIPLSKRDAGSSAAAMRSPSIT	
		OLVDFRFRKPTGSSATACNTPCRDYACGTGGTIDKIDYICGMFTATWISPYVEQIT	
		DDTGTGATVHGTVMKDIYALINSFETADLNLNSLNLKRMKMLAVDIYVHYAMNGA	
		GSSVAVSNYPNOOSYFHDYCLITNDDQNVDECMWGDNTVLSIPDRTEDSVSS	
		FLNWAEVLVSNYSIDGLRIDSAKHVDESFYSPFAAGVYLLGEVYDGPAYTCPOYQ	
		YMSGATNPLPLYPMILREFQGTNSVDELNAMISSLESDCKDITLLGNITENHDDPRL	
		SYTSNPALIKNAIAFNLMSDGIPIITYYEGEGYSSSDPNRRLMLTSGYSTSNVY	
		LIISSNORNAIYKDKSYTTWSDVYASGHVIALQRCADQDRIYSVFNNLSSGSG	
		TWTSIKTYSGGEKYVDVLTCTSTANSDTLTIVSISGAPRIIAPASLIANSIGNF	
BASE COUNT	654 a	375 c	379 g
ORIGIN		735 t	
Query Match	23.6%	Score 409;	DB 8; Length 2143;
Best Local Similarity	58.4%;	Pred. No. 9,8e-86;	
Matches 715; Conservative	0;	Mismatches 510;	Indels 0; Gaps 0;
QY	121	ctgcgcagctggcagatcgcaatcatattcttccttcacgcagatcgatcttgcagagcg	180
Db	337	CTGCAGCTTGGCGGTTCTGATCTATCTATCAACACTGTATTCCAGATAGATTGGCAGACAG	396
QY	181	atgggttcgcacgactgcgacttgtaactcgcgatacgaataactgtgtggaacatggc	240

Db 397 ACGGATCGAGCTTACGCTACTTGTAAATCTGACATGAGATTAATCTGTGGGGGACTTTCC 456
QY 241 agggacatcatcgacacattggaactatataccaggaatggtcttaagacatctgatala 300
Db 457 AAGGTATTATTGACAAATTTGGATTACATCCAAAGTATGGGTTCATGCTATTTTGATT 516
QY 301 ccccgcttaacagccagctgccccagaccacccatagatgagatgctccctcctgact 360
Db 517 CTCGAGTGTGTGAACAATTCCTGATGATACCTGTATGTTATGCTTACCAAGGCTATT 576
QY 361 ggcagcagatatatactctctgaacaaactaagcagctgcagatgacttgaagcgc 420
Db 577 GGATGAAGATATTACGCTAATTAATTCAAATTTTGTACTCCGATGACTGGAAGATC 636
QY 421 tctctcgccctcatgagagggatglatctatgctgagtgtgtgtaaccata 480
Db 637 TTTCAAATGATTTGCAATGAAGAAATATGAAGCTTATGGTATGTTACTTAACCAT 696
QY 481 tgggctatgataggagcgggtagctcagtcagttacagtggtttaaccgttcagttccc 540
Db 697 ATGCTTGAATGATGGCGGAGAGAGTGCTTACTCCAACTACAAATCCATTCACCAAC 756
QY 541 aagactacttcaacccgttctgttcatcaaaaactataagatcagactcagttgag 600
Db 757 AATCCTACTTCCACGATTTATTGTTAATTACAAATTACGATGATCAACCAATGTTGAAG 816
QY 601 attgctgcttagagataacacgtctcctctgctgctgctgctgctgctgctgctgctg 660
Db 817 ATTGCTGGGAAGGCGATTAATACCTTAAATACCATCTTCCAGATCTTCCAGATG 876
QY 661 tcaagaatgaatgtaagactggtggtggtatggtatgtaactactcattgaagcgc 720
Db 877 TTAGCTCTATTTCATCTGCTGGTGGTGGAGTACTTCTAATTAATTAATGATGTT 936
QY 721 tccgtatcgacacagtaaaacacgtccagaaggaactctgcccgggtacaaagaagcgc 780
Db 937 TAAGGATTCACAGCTAAGCATGTGATGAATTCATTCACCATTCACCAAGTGGCTG 996
QY 781 caagcgctgactgctcggcggaggtctcgaagcgtgactcggcctgactgctcctacc 840
Db 997 CAGGTCTTATCTCTTGGAGAAATTTATGACGGTATCCAGCTTACACTTCCCATACC 1056
QY 841 aagaagctatgagcggcgtactgaactatccattactatccactcctcaacgcctca 900
Db 1057 AAACATATATGTCAGGGTACTTAATCTTGTACTATCCAAATGTAAGATCTTTC 1116
QY 901 agtcaacctccggcagcatggaagcactctacaacatgatacaacgcgtcaaatccgact 960
Db 1117 AAGGTACTTCTAATCTGTCGATGAATTAATGCTATGATTTCAAGTTTGAAGATGATT 1176
QY 961 gtccagactcaaacactcctcgggacatctcgtcgaagacacagcaaccacggcttcgt 1020
Db 1177 GTTAGGATTAATTTATTTGGGTAATTTTCATGTAAGAAACCATGATCAACCAAGATTACAT 1236
QY 1021 cttaacacaaagacatagccctcgcgaagaagcgtcgcagcatlcatcctcaagcgc 1080
Db 1237 CTTATCTTCTGATAGTGCTTAATCAAAAATGCAATTCGTTAATTATATGTCAGATG 1296
QY 1081 gaatcccatctcatagccggcgaagaacacacactacgcggcggaagcccccgcga 1140
Db 1297 GTATTCCAAATTAATTAAGGTCAGGAACAAGGTTACGTTAGCTCCGATCCCAACA 1356
QY 1141 accggaagcaaacctgctcgcgtacacgcagacagcagcgtgtacaagttacttg 1200
Db 1357 ACAGAGAAACATTAATGCTTATCTGTTACAGACACTAGTATGTTACTRCAAACTTATCT 1416
QY 1201 cctcgcggaacgcaatccggaactatgcccattagaagaatacagatctgtaacctaca 1260
Db 1417 CTTCAATTAATTAAGAAACCAAGCATTTATTAAGGATGCAAAATACACTACTTATT 1476
QY 1261 agaactgcccactctacaagaagacaaagcatcgccatgcgaagggcacagatggct 1320

Db 1477 GGAGTATGTTATACGCTTCAGGTCATGTTATGCTTCAAGAGAGTGACAGCACC 1536
QY 1321 cgcagatcgtactatctgtccaa 1345
Db 1537 AAAGATTTGTTCTGCTTTAACA 1561

Search completed: June 13, 2002, 18:23:34
Job time: 16672 sec

Fri Jun 14 10:36:35 2002

us-09-710-339-1.rge

Page 25

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 13, 2002, 08:54:29 ; Search time 57.95 Seconds
(without alignments)
4.636 Million cell updates/sec

Title: US-09-710-339-2_COPY_150_160
Perfect score: 57
Sequence: 1 SSVDSYVKPF 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PC10S.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	478	1	US-08-720-899-7 Sequence 7, Appli
2	57	100.0	478	1	US-08-459-610-7 Sequence 7, Appli
3	57	100.0	478	1	US-08-343-804-7 Sequence 7, Appli
4	57	100.0	478	2	US-08-339-715A-2 Sequence 2, Appli
5	57	100.0	478	2	US-08-600-908A-10 Sequence 10, Appli
6	57	100.0	478	3	US-08-683-838A-10 Sequence 10, Appli
7	57	100.0	478	4	US-09-182-859-7 Sequence 7, Appli
8	37	64.9	468	1	US-08-204-656B-2 Sequence 2, Appli
9	37	64.9	468	1	US-08-204-656B-4 Sequence 4, Appli
10	37	64.9	468	1	US-08-204-656B-6 Sequence 6, Appli
11	37	64.9	468	1	US-08-204-656B-8 Sequence 8, Appli
12	37	64.9	468	1	US-08-470-702-6 Sequence 6, Appli
13	37	64.9	468	1	US-08-470-702-7 Sequence 7, Appli
14	37	64.9	468	1	US-08-470-702-8 Sequence 8, Appli
15	37	64.9	468	1	US-08-470-702-9 Sequence 9, Appli
16	37	64.9	468	1	US-08-467-831-6 Sequence 6, Appli
17	37	64.9	468	1	US-08-467-831-7 Sequence 7, Appli
18	37	64.9	468	1	US-08-467-831-8 Sequence 8, Appli
19	37	64.9	468	1	US-08-467-831-9 Sequence 9, Appli
20	35	61.4	349	4	US-09-105-390-48 Sequence 48, Appli
21	35	61.4	377	4	US-09-105-390-64 Sequence 64, Appli
22	35	61.4	497	4	US-09-058-947A-4 Sequence 4, Appli
23	35	61.4	500	4	US-08-868-373-12 Sequence 12, Appli
24	33	57.9	1507	6	5268270-2 Patent No. 5268270
25	33	57.9	1541	4	US-08-296-791-3 Sequence 3, Appli
26	33	57.9	1541	5	PCT-US95-10661A-3 Sequence 3, Appli
27	33	57.9	1545	4	US-08-296-791-4 Sequence 4, Appli

28	33	57.9	1545	5	PCT-US95-10661A-4 Sequence 4, Appli
29	33	57.9	1702	4	US-08-296-791-5 Sequence 5, Appli
30	33	57.9	1702	5	PCT-US95-10661A-5 Sequence 5, Appli
31	33	57.9	1848	4	US-08-296-791-6 Sequence 6, Appli
32	33	57.9	1848	5	PCT-US95-10661A-6 Sequence 6, Appli
33	32	56.1	172	1	US-08-487-001A-51 Sequence 51, Appli
34	32	56.1	172	2	US-08-630-822A-51 Sequence 51, Appli
35	32	56.1	172	2	US-09-005-069-51 Sequence 51, Appli
36	32	56.1	182	2	US-08-874-832-14 Sequence 14, Appli
37	32	56.1	182	2	US-09-097-233-14 Sequence 14, Appli
38	32	56.1	213	2	US-08-874-832-5 Sequence 5, Appli
39	32	56.1	213	3	US-09-097-233-5 Sequence 5, Appli
40	32	56.1	425	4	US-09-247-155-99 Sequence 99, Appli
41	32	56.1	510	1	US-08-255-670A-2 Sequence 2, Appli
42	32	56.1	675	1	US-08-317-522A-9 Sequence 9, Appli
43	32	56.1	675	1	US-08-439-818A-9 Sequence 9, Appli
44	32	56.1	675	2	US-08-751-965-9 Sequence 9, Appli
45	32	56.1	675	2	US-08-738-975-9 Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-08-720-899-7
Sequence 7, Application US/08720899
Patent No. 5753460
GENERAL INFORMATION:
APPLICANT: Bisgaard-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
APPLICANT: Svendsen, Allan
APPLICANT: Thellersen, Marianne
APPLICANT: Van der Zee, Pia
TITLE OF INVENTION: AMYLASE VARIANTS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 57534600 No. 5753460disk of No. 5753460th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,899
FILING DATE: 10-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/343,804
FILING DATE: 22-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4054, 214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-720-899-7
Query Match 100.0%; Score 57; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11
1111111111
Db 130 SSVDSVFKPF 140

RESULT 2
US-08-459-610-7
; Sequence 7, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58010430 No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowmey Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-459-610-7

Query Match 100.0%; Score 57; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11
1111111111
Db 130 SSVDSVFKPF 140

RESULT 3
US-08-343-804-7
; Sequence 7, Application US/08343804
; Patent No. 5830837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS

NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58308370 No. 5830837disk of No. 5830837th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,804
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowmey Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-343-804-7

Query Match 100.0%; Score 57; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11
1111111111
Db 130 SSVDSVFKPF 140

RESULT 4
US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michio
; APPLICANT: Takata, Hiroki
; APPLICANT: Okada, Shigetaka
; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POPHAM, HAIK, SCHNOBIRCH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,715A
; FILING DATE: 14-NOV-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Melser, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335.009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 100.0%; Score 57; DB 2: Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
|||||
Db 130 SSVDSYVFKPF 140

RESULT 5
US-08-600-908A-10
Sequence 10, Application US/08600908A
Patent No. 5989169
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borcherdt, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59891690 No. 59891690th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,908A
FILING DATE: 13-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-600-908A-10

Query Match 100.0%; Score 57; DB 2: Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
|||||
Db 130 SSVDSYVFKPF 140

RESULT 6
US-08-683-838A-10
Sequence 10, Application US/08683838A
Patent No. 6022724
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borcherdt, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60227240 No. 6022724th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,838A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/600,908
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-683-838A-10

Query Match 100.0%; Score 57; DB 3: Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
|||||
Db 130 SSVDSYVFKPF 140

RESULT 7
US-09-182-859-7
Sequence 7, Application US/09182859
Patent No. 6143708
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Borcherdt, Torben
APPLICANT: Bisgard-Frantzen, Henrik

;; TITLE OF INVENTION: Alpha-Amylase Mutants
;; FILE REFERENCE: 4796.204-US
;; CURRENT APPLICATION NUMBER: US/09/182,859
;; CURRENT FILING DATE: 1998-10-29
;; EARLIER APPLICATION NUMBER: 0515/96
;; EARLIER FILING DATE: 1996-04-30
;; EARLIER APPLICATION NUMBER: 0712/96
;; EARLIER FILING DATE: 1996-06-28
;; EARLIER APPLICATION NUMBER: 0775/96
;; EARLIER FILING DATE: 1996-07-11
;; EARLIER APPLICATION NUMBER: 1263/96
;; EARLIER FILING DATE: 1996-11-08
;; NUMBER OF SEQ ID NOS: 37
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 7
;; LENGTH: 478
;; TYPE: PRT
;; ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match 100.0%; Score 57; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDYSEVFKPF 11
Db 130 SVDYSEVFKPF 140

RESULT 8
US-08-204-656B-2
; Sequence 2, Application US/08204656B
; Patent No. 5538882
; GENERAL INFORMATION:
; APPLICANT: Matsui, Ikuro
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Miyairi, Sachio
; APPLICANT: Honda, Koichi
; TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
; TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
; TITLE OF INVENTION: Oligosaccharide Using The Enzyme
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,656B
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-204-656B-2

Query Match 64.9%; Score 37; DB 1; Length 468;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SVDYSEVFKPF 11
Db 132 SVDYSEVFKPF 141

RESULT 9
US-08-204-656B-4
; Sequence 4, Application US/08204656B
; Patent No. 5538882
; GENERAL INFORMATION:
; APPLICANT: Matsui, Ikuro
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Miyairi, Sachio
; APPLICANT: Honda, Koichi
; TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
; TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
; TITLE OF INVENTION: Oligosaccharide Using The Enzyme
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,656B
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-204-656B-4

Query Match 64.9%; Score 37; DB 1; Length 468;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SVDYSEVFKPF 11
Db 132 SVDYSEVFKPF 141

RESULT 10
US-08-204-656B-6
; Sequence 6, Application US/08204656B
; Patent No. 5538882
; GENERAL INFORMATION:
; APPLICANT: Matsui, Ikuro

APPLICANT: Ishikawa, Kazuhiko
APPLICANT: Miyairi, Sachio
TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
TITLE OF INVENTION: Oligosaccharide Using The Enzyme
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-204-656B-6

Query Match 64.9%; Score 37; DB 1; Length 468;
Best local similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 SYDYSVFKPF 11
: : : : :
Db 132 SIDYSEYTF 141

RESULT 11
US-08-204-656B-8
Sequence 8, Application US/08204656B
Patent No. 5538882
GENERAL INFORMATION:
APPLICANT: Matsui, Ikuro
APPLICANT: Ishikawa, Kazuhiko
APPLICANT: Miyairi, Sachio
APPLICANT: Honda, Koichi
TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
TITLE OF INVENTION: Oligosaccharide Using The Enzyme
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-204-656B-8

Query Match 64.9%; Score 37; DB 1; Length 468;
Best local similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 SYDYSVFKPF 11
: : : : :
Db 132 SIDYSEYTF 141

RESULT 12
US-08-470-702-6
Sequence 6, Application US/08470702
Patent No. 5631149
GENERAL INFORMATION:
APPLICANT: MATSUI, IKURO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,702
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 468 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-470-702-6

Query Match 64.9%; Score 37; DB 1; Length 468;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 SYDYSVEKPF 11
1:111:11
Db 132 SIDYSEYTFP 141

RESULT 13
US-08-470-702-7
; Sequence 7, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,702
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,656
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-470-702-7

Query Match 64.9%; Score 37; DB 1; Length 468;
Best Local Similarity 60.0%; Pred. No. 27;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 2 SYDYSVEKPF 11
1:111:11
Db 132 SIDYSEYTFP 141

RESULT 14
US-08-470-702-8
; Sequence 8, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,702
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,656
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-470-702-8

Query Match 64.9%; Score 37; DB 1; Length 468;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 SYDYSVEKPF 11
1:111:11
Db 132 SIDYSEYTFP 141

RESULT 15
US-08-470-702-9
; Sequence 9, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:

APPLICANT: MATSUI, IKUO
 APPLICANT: ISHIKAWA, KAZUHIKO
 APPLICANT: MIYAIRI, SACHIO
 APPLICANT: HONDA, KOICHI
 TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
 TITLE OF INVENTION: VARIANT GENE OF THE ENZYME FOR PRODUCING
 TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
 STREET: 810 Gatehouse Road, Suite 500 East
 CITY: Falls Church
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22042
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/470,702
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/204,656
 FILING DATE: 02-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: WEINER, MARC S.
 REGISTRATION NUMBER: 32,181
 REFERENCE/DOCKET NUMBER: 234-252P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 205-8000
 TELEFAX: (703) 205-8050
 TELEX: 248345
 INFORMATION FOR SEQ. ID NO. 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 468 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-470-702-9

Query Match 64.9%; Score 37; DB 1; Length 468;
 Best local Similarity 60.0%; Pred. No. 27;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 SVDSVFKPF 11
 1:111:11
 DB 132 SIDISEYTPF 141

Search completed: June 13, 2002, 08:54:30
 Job time: 315 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 13, 2002, 09:11:07 ; Search time 125.81 seconds
(without alignments)
15.126 Million cell updates/sec

Title: US-09-710-339-2_COPY_150_160

Perfect score: 57

Sequence: 1 SSVDYSVFKPF 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	499	3	Q96TH4
2	45	78.9	490	3	Q9UV07
3	44	77.2	640	3	O13296
4	39	68.4	923	6	O97745
5	39	68.4	945	6	O77589
6	39	68.4	964	6	O97744
7	39	68.4	964	6	Q9T001
8	39	68.4	964	6	Q9T001
9	39	68.4	978	6	O9XS93
10	39	68.4	4116	5	O96958
11	39	68.4	5322	5	O9XYD2
12	39	68.4	5322	5	O9VLT5
13	38	66.7	350	10	Q9FGE2
14	38	66.7	425	16	O9KRR9
15	38	66.7	430	3	O13699
16	38	66.7	623	3	Q9UV09

17	38	66.7	1869	3	O93809	O93809 magnaporthe
18	38	66.7	2160	5	O9U5P3	O9U5P3 halictis tu
19	38	66.7	2500	5	O9NDV3	O9NDV3 halictis tu
20	37	64.9	3419	5	O9GP18	O9GP18 halictis tu
21	37	64.9	245	8	O02665	O02665 podospora a
22	37	64.9	790	4	O96MR9	O96MR9 homo sapien
23	37	64.9	4820	5	O9VAV5	O9VAV5 drosophila
24	36	63.2	112	10	O9LZU8	O9LZU8 arabidopsis
25	36	63.2	173	5	O9XNM1	O9XNM1 caenorhabdi
26	36	63.2	257	16	O98IS0	O98IS0 rhizobium 1
27	36	63.2	300	16	O9CGG9	O9CGG9 lactococcus
28	36	63.2	353	5	O9XVC7	O9XVC7 caenorhabdi
29	36	63.2	409	10	O9SCU1	O9SCU1 arabidopsis
30	36	63.2	491	3	O13996	O13996 schizosach
31	36	63.2	499	3	O9UT42	O9UT42 schizosach
32	36	63.2	541	2	O9ZNC7	O9ZNC7 clostridium
33	36	63.2	579	11	O9WV10	O9WV10 mus muscui
34	36	63.2	1301	4	O9GCR7	O9GCR7 homo sapien
35	36	63.2	1329	4	O9HCR3	O9HCR3 homo sapien
36	36	63.2	2692	5	O21547	O21547 caenorhabdi
37	35.5	62.3	690	16	O98IT8	O98IT8 rhizobium 1
38	35	61.4	143	8	O9XNE1	O9XNE1 erlichynchu
39	35	61.4	363	1	O9UWQ2	O9UWQ2 pyrococcus
40	35	61.4	377	10	O9ZNY6	O9ZNY6 oryza sativ
41	35	61.4	428	5	O24024	O24024 drosophila
42	35	61.4	428	5	O9VFR3	O9VFR3 drosophila
43	35	61.4	483	4	O9H798	O9H798 homo sapien
44	35	61.4	492	10	O9C6L5	O9C6L5 arabidopsis
45	35	61.4	497	10	O9XPF3	O9XPF3 arabidopsis

ALIGNMENTS

RESULT	ID	Q96TH4	PRELIMINARY:	PRT:	499 AA.
AC	O96TH4				
DT	01-DEC-2001	(TREMBLrel. 19, Created)			
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)			
DE	TAKA-AMYLASE A (EC 3.2.1.1)	(ALPHA-AMYLASE).			
GN	AMYA OR AMY1.				
OS	Aspergillus oryzae, and				
OS	Aspergillus flavus.				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				
OX	NCBI_TaxID=5062, 5059;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=A.oryzae; STRAIN=RI840;				
RX	MEDLINE=20289310; PubMed=10830498;				
RA	Gomi K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N.,				
RA	Imura Y.;				
RT	"Molecular cloning and characterization of a transcriptional activator				
RT	gene, amyR, involved in the amylolytic gene expression in Aspergillus				
RT	oryzae.";				
RL	Biosci. Biotechnol. Biochem. 64:816-827(2000).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=A.flavus; STRAIN=86-10D;				
RA	Fakhoury A.M., Woloshuk C.P.;				
RT	"Amy1, the alpha-amylase gene of Aspergillus flavus: Involvement in				
RT	atactoxin biosynthesis in maize kernels.";				
RL	Phytopathology 89:908-914(1999).				
DR	EMBL; AB021876; BAA95703.1; -;				
DR	EMBL; AF139925; AAF14264.1; -;				
SQ	SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;				

Query Match 100.0%; Score 57; DB 3; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0091;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
 |||||
 Db 151 SSVDSYVFKPF 161

RESULT 2
 Q9UV07 PRELIMINARY; PRT; 490 AA.

AC 09UV07;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ALPHA-AMYLASE AMYA.
 GN AMYA.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OC NCBI_TaxID=5072;

RM [1]
 RP SEQUENCE FROM N.A.
 RA Boase N.A., Murphy R.M., Kelly J.M.;
 RT "An amylase cluster in Aspergillus nidulans."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF208225; AAF17103.1; -.
 DR HSSP; P10529; 7TAA.
 DR InterPro; IPR000461; Alpha-amylase.
 DR Pfam; PF00128; alpha-amylase; 1.
 DR Pfam; PF02806; alpha-amylase.C; 1.
 DR PRINTS; PR00110; ALPHAMYLASE.
 SO SEQUENCE 490 AA; 5429 MW; A891C4ACEAE5305 CRC64;

Query Match 78.9%; Score 45; DB 3; Length 490;
 Best Local Similarity 72.7%; Pred. No. 1.9;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
 :|||||
 Db 143 ASVDYSAENPF 153

RESULT 3
 O13296 PRELIMINARY; PRT; 640 AA.

AC 013296;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ACID-STABLE ALPHA-AMYLASE.
 OS Aspergillus kawachii.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OC NCBI_TaxID=40384;

RM [1]
 RP SEQUENCE FROM N.A.
 RA Kaneko A.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 [2]

RP SEQUENCE FROM N.A.
 RA Kaneko A., Sudo S., Sakamoto Y., Tamura G., Ishikawa T., Ohba T.;
 RT "Molecular-cloning and determination of the nucleotide-sequence of a
 RT gene encoding an acid-stable alpha-amylase from Aspergillus-
 RT kawachii."
 RL J. Ferment. Bioeng. 81:292-298(1996).
 DR EMBL; AB008370; BAA22993.1; -.
 DR HSSP; P56271; 2AAA.
 DR InterPro; IPR000461; Alpha-amylase.
 DR InterPro; IPR002044; CBD_4.
 DR Pfam; PF00128; alpha-amylase; 1.
 DR Pfam; PF00686; CBD_4; 1.
 DR ProDom; PD001568; CBD_4; 1.
 SO SEQUENCE 640 AA; 69507 MW; 676BF9D0236720DD CRC64;

Query Match 77.2%; Score 44; DB 3; Length 640;
 Best Local Similarity 72.7%; Pred. No. 3.9;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSVDSYVFKPF 11
 :|||||
 Db 151 NDVDSYVFDPF 161

RESULT 4
 O97745 PRELIMINARY; PRT; 923 AA.

AC 097745;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE MAST/STEM CELL GROWTH FACTOR RECEPTOR (FRAGMENT).
 GN KIT2*0201.

OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OC NCBI_TaxID=9823;

RM [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
 RA Marklund S.,
 RL Thesis (1997), Department of Animal Breeding and Genetics,
 RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
 RX MEDLINE=98391767; PubMed=9724328;
 RA Marklund S., Kijas J., Rodriguez-Martinez H., Roenstrand L., Funa K.,
 RA Moller M., Edfors-Lilja I., Andersson L.;
 RT "Molecular evolution of the dominant white phenotype in the domestic
 RT pig."
 RL Genome Res. 8:826-833(1998).

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.

DR EMBL; AJ223231; CA11199.1; -.

DR HSSP; P11362; 1FGK.

DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR003600; Ig_Like.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR001824; Receptor_tyr_kin_III.

DR InterPro; IPR001245; Tyr_pkinase.

DR Pfam; PF00047; Ig_2.

DR Pfam; PF00069; pkinase; 3.

DR PRINTS; PR00109; TYRKINASE.

DR SMART; SM00409; IG_1.

DR SMART; SM00410; IG_Like; 1.

DR SMART; SM00219; TYRK; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ARP; 1.

DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.

DR APP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
 KW Receptor; Transferase; Transmembrane; Tyrosine-protein kinase.

FT NON_TER 923
 SO SEQUENCE 923 AA; 103780 MW; FDCD566F5D86B9F CRC64;

Query Match 68.4%; Score 39; DB 6; Length 923;
 Best Local Similarity 77.8%; Pred. No. 53;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSVDSYVFK 9
 |||||
 Db 476 SSIDYSAFK 484

```

RESULT 5
ID 077589 PRELIMINARY; PRT: 945 AA.
AC 077589;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TYROSINE KINASE RECEPTOR HOMOLOG (FRAGMENT).
GN KIT.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Lear T.L., Adams M.H., McDowell K.J., Ponce de Leon F.A.,
RA Sullivan N.D., Coogee U., Ferguson E., Chambers T.M., Bailey E.;
RT "Chromosomal location of genes for ESR, ETS2, GOR2, KIT, MXI and PGR
RT in the horse, Equus caballus.";
RL Proc. Xth Amer. Coll. Dom. Anim. Cytogenet. Gene Mapp. 0:0-0(1997).
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Rieder S., Checa-Cortes M.L., Joerg H., Stranzinger G.;
RT "An Equine sequence homologous to the c-KIT-proto-oncogene mapped to
RT chromosome 3q13-21.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: AF055037; AAC23611.1; -.
DR HSSP: P11362; IFGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; Receptor_tyr_kin_III.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00069; pkinase; 2.
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SM00409; IG; 2.
DR SMART: SM00219; TYKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
FT NON_TER 945
FT SEQUENCE 945 AA; 106143 MW; 7DB424A89DE5C340 CRC64;

Query Match 68.4%; Score 39; DB 6; Length 945;
Best Local Similarity 77.8%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSVDSYVK 9
DB 476 SSIDYSAFK 484

```

```

RESULT 6
O97744 PRELIMINARY; PRT: 964 AA.
AC O97744;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MAST/STEM CELL GROWTH FACTOR RECEPTOR (FRAGMENT).
GN KIT1*0101.

```

```

OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=HAMPSHIRE BREED; TISSUE=LEUKOCYTE;
RX MEDLINE=98391767; PubMed=9724328;
RA Marklund S., Kijas J., Rodriguez-Martinez H., Reenstrand L., Funä K.,
RA Woller M., Editors-Lilla I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
RT pig.";
RL Genome Res. 8:826-833(1998).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN=HAMPSHIRE BREED; TISSUE=LEUKOCYTE;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: AJ223228; CA01196.1; -.
DR HSSP: P11362; IFGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; Receptor_tyr_kin_III.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00069; pkinase; 2.
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SM00409; IG; 1.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00219; TYKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DR ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
KW Receptor; Transferase; Transmembrane; Tyrosine-protein kinase.
FT NON_TER 964
FT SEQUENCE 964 AA; 108287 MW; 90E72EDFAB1358B7 CRC64;

Query Match 68.4%; Score 39; DB 6; Length 964;
Best Local Similarity 77.8%; Pred. No. 55;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSVDSYVK 9
DB 476 SSIDYSAFK 484

```

```

RESULT 7
O97001 PRELIMINARY; PRT: 964 AA.
AC O97001;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MAST/STEM CELL GROWTH FACTOR RECEPTOR (FRAGMENT).
GN KIT1*0201.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,

```

RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
 RX MEDLINE=98391767; PubMed=9724328;
 RA Marklund S., Kijas J., Rodriguez-Martinez H., Roenstrand L., Funa K.,
 RA Moller M., Edfors-Lilja I., Andersson L.;
 RT "Molecular evolution of the dominant white phenotype in the domestic
 pig."
 RL Genome Res. 8:826-833(1998).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 DR EMBL: AJ223229; CAI1197.1; -.
 DR HSSP: P11362; 1FGK
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003600; Ig_like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR001824; Receptor_tyr_kin_III.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00047; Ig; 2.
 DR Pfam: PF00069; pkinase; 2.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00409; IG; 1.
 DR SMART: SM00219; TYRKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
 KW ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
 KW Receptor; Transferase; Transmembrane; Tyrosine-protein kinase.
 FT NON_TER 964 964
 SQ SEQUENCE 964 AA; 108315 MW; 996C3C46201356A8 CRC64;

Query Match 68.4%; Score 39; DB 6; Length 964;
 Best Local Similarity 77.8%; Pred. No. 55;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SSVDSVFK 9
 DB 476 SSIDYSAFK 484
 RESULT 8
 O9TQO0 PRELIMINARY; PRT; 964 AA.
 AC O9TQO0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MAST/STEM CELL GROWTH FACTOR RECEPTOR (FRAGMENT).
 GN K11*0202.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
 RA Marklund S.;
 RL Thesis (1997), Department of Animal Breeding and Genetics,
 RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
 RX MEDLINE=98391767; PubMed=9724328;
 RA Marklund S., Kijas J., Rodriguez-Martinez H., Roenstrand L., Funa K.,
 RA Moller M., Edfors-Lilja I., Andersson L.;
 RT "Molecular evolution of the dominant white phenotype in the domestic
 pig.";

RL Genome Res. 8:826-833(1998).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 DR EMBL: AJ223230; CAI1198.1; -.
 DR HSSP: P11362; 1FGK
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003600; Ig_like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR001824; Receptor_tyr_kin_III.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00047; Ig; 2.
 DR Pfam: PF00069; pkinase; 2.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00409; IG; 1.
 DR SMART: SM00219; TYRKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
 KW ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
 KW Receptor; Transferase; Transmembrane; Tyrosine-protein kinase.
 FT NON_TER 964 964
 SQ SEQUENCE 964 AA; 108287 MW; 1B21292A962E9191 CRC64;

Query Match 68.4%; Score 39; DB 6; Length 964;
 Best Local Similarity 77.8%; Pred. No. 55;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SSVDSVFK 9
 DB 476 SSIDYSAFK 484
 RESULT 9
 O9XS93 PRELIMINARY; PRT; 978 AA.
 AC O9XS93;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE KIT.
 GN C-KIT.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=CEREBELLUM;
 RA London C.A., Galli S.J., Yuuki T., Hu Z.-Q., Helfand S.C.,
 RA Geissler E.N.;
 RT "Spontaneous canine mast cell tumors express tandem duplications in
 the proto-oncogene c-kit.";
 RL Exp. Hematol. 0:0-0(1999).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 DR EMBL: AF099030; AAD28369.1; -.
 DR HSSP: P11362; 1FGK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003600; Ig_like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR001824; Receptor_tyr_kin_III.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00047; Ig; 2.
 DR Pfam: PF00069; pkinase; 2.
 DR SMART: SM00409; IG; 2.
 DR SMART: SM00219; TYRKC; 1.


```
DR SMART: SM00219; TYRKC: 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP: 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM: 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR: 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_TIR: 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
SQ Receptor; Transferrase; Transmembrane; Tyrosine-protein kinase.
SEQUENCE 978 AA; 109651 MW; 5AC31E5AC4E9910F CRC64;

Query Match
Best Local Similarity 68.4%; Score 39; DB 6; Length 978;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSVDYSVFK 9
   ||:|||||
Db 478 SSIDYSAFK 486

RESULT 10
ID 096958 PRELIMINARY; PRT; 4116 AA.
AC 096958:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CALO PROTEIN (FRAGMENT).
GN POE OR CALO OR CG14472.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RETINAL;
RX MEDLINE=99030403; PubMed=9813038;
RA Xu X.Z.S., Wes P.D., Chen H., Li H.S., Yu M., Morgan S., Liu Y.,
RA Montell C.;
RT "Retinal targets for calmodulin include proteins implicated in
RT synaptic transmission.";
RL J. Biol. Chem. 273:31297-31297(1998).
DR EMBL: Y17920; CAAT6940.1; -.
DR FlyBase: FBgn0011230; Poe.
DR InterPro: IPR002106; AA_TRNA_Ligase_II.
DR InterPro: IPR003126; Znf_UBR1.
DR Pfam: PF02207; Zf-UBR1.1.
DR SMART: SM00396; Znf_UBR1.1.
DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 4116 AA; 457563 MW; D3DFB691E7EF3F36 CRC64;

Query Match
Best Local Similarity 68.4%; Score 39; DB 5; Length 4116;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 DYSVFKPF 11
   ||:|||||
Db 4013 DYTIVKPF 4020

RESULT 11
O9XYD2 PRELIMINARY; PRT; 5322 AA.
AC 09XYD2:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUSHOVER.
GN POE OR PUSH OR CG14472.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
```

```
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Richards S., Hekmat-Scafe D.S., Hurd D.D., Caprette D.R., Saxton W.M.,
RA Carlson J.R., Stern M.J.;
RT "Hyperplastic glial cells surrounding abdominal nerves in pushover and
RT inbredated pushover mutants of Drosophila melanogaster.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF096897; AAD20450.1; -.
DR FlyBase: FBgn0011230; poe.
DR InterPro: IPR002106; AA_TRNA_Ligase_II.
DR InterPro: IPR003126; Znf_UBR1.
DR Pfam: PF02207; Zf-UBR1.1.
DR SMART: SM00396; Znf_UBR1.1.
DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
FT VARIANT 728 728 Q -> *.
FT VARIANT 883 883 Q -> *.
SQ SEQUENCE 5322 AA; 590679 MW; ABP4C4BC7A9CEBFE CRC64;

Query Match
Best Local Similarity 68.4%; Score 39; DB 5; Length 5322;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 DYSVFKPF 11
   ||:|||||
Db 5219 DYTIVKPF 5226

RESULT 12
O9VLT5 PRELIMINARY; PRT; 5322 AA.
AC 09VLT5:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CG14472 PROTEIN.
GN POE OR CG14472.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazek R.G., Chame M., Pfeiffer B.D.,
RA Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Aghayani A., An H.-J., Andrews-Plankkoch C., Baldwin D.,
RA Balieu R.M., Basu A., Baxendale J., Bayraktaroglu U., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dabike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
```

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclele J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinelt K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Styrcas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao S., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Glibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "the genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003619; AAF52598.1; -.
 DR FLYBase: FBgn0011230; poe.
 DR InterPro: IPR002106; AA_TRNA_Ligase_II.
 DR InterPro: IPR003126; znf-UBR1.
 DR Pfam: PF02207; zf-UBR1.1.
 DR SMART: SM00396; znf-UBR1.1.
 DR PROSITE: PS00339; AA_TRNA_LIGASE_II.2; UNKNOWN.1.
 DR SEQUENCE 5322 AA; 590696 MW; 3890824E00F2DB50 CRC64;

Query Match 68.4%; Score 39; DB 5; Length 5322;
 Best Local Similarity 75.0%; Pred. No. 3.4e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 4 DYSVFKPF 11
 DB 5219 DYTYYKPF 5226

RESULT 13
 Q9FGE2 PRELIMINARY; PRT; 350 AA.
 AC Q9FGE2;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE GBI/AD2930.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucotids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
 RL Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB025641; BAB08931.1; -.
 SEQUENCE 350 AA; 39471 MW; C33011C2DB86C2DB CRC64;

Query Match 66.7%; Score 38; DB 10; Length 350;
 Best Local Similarity 54.5%; Pred. No. 29;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SSVDSVFKPF 11
 DB 329 SSMDYKSYKPF 339

RESULT 14
 ID Q9KRR9 PRELIMINARY; PRT; 425 AA.
 AC Q9KRR9;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE HYPOTHETICAL PROTEIN VC1567.
 GN VC1567.
 OS *Vibrio cholerae*.
 OC Bacteria; Proteobacteria; gamma subdivision: Vibrionaceae; *Vibrio*.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodsan R.D., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.";
 RL Nature 406:477-483(2000).
 DR EMBL: AE004234; AAF94721.1; -.
 DR TIGR: VC1567; -.
 DR InterPro: IPR003838; DUF214.
 DR Pfam: PF02687; DUF214; 1.
 DR Complete proteome.
 DR SEQUENCE 425 AA; 46328 MW; 982D2C2DBAFA9EE7 CRC64;

Query Match 66.7%; Score 38; DB 16; Length 425;
 Best Local Similarity 63.6%; Pred. No. 36;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SSVDSVFKPF 11
 DB 209 SPIDYSGFTPF 219

RESULT 15
 ID O13699 PRELIMINARY; PRT; 430 AA.
 AC O13699;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE HYPOTHETICAL 50.4 KDA PROTEIN C11G7.06C IN CHROMOSOME I.
 GN SPAC11G7.06C.
 OS Schizosaccharomyces pombe (fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-972;
 RA McLean J., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
 RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
 CC -I- SIMILARITY: SOME, TO THE C-TERMINAL OF S.POMBE SPAC4G9.07.
 DR EMBL: Z99161; CAB16211.1; -.
 KW Hypothetical protein.
 SEQUENCE 430 AA; 50433 MW; 6FE023D9545B286A CRC64;

Query Match 66.7%; Score 38; DB 3; Length 430;
 Best Local Similarity 70.0%; Pred. No. 36;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 SVDYSVFKPF 11
 DB 21 SIDYSQFKAF 30

Search completed: June 13, 2002, 09:11:10
 Job time: 1055 sec

Fri Jun 14 10:36:43 2002

us-09-710-339-2_copy_150_160.rspt

Page 7

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 13, 2002, 09:11:49 ; Search time 38.1 Seconds

(without alignments)
11.179 Million cell updates/sec

Title: US-09-710-339-2_COPY_150_160

Perfect score: 57

Sequence: 1 SSVDTSVKPF 11

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 105224 segs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	100.0	498	1 AMYA_ASPAW	002905 aspergillus
2	57	100.0	499	1 AMYA_ASPOR	P10529 aspergillus
3	57	100.0	499	1 AMYA_ASPAW	Q02906 aspergillus
4	57	100.0	499	1 AMYA_ASPSH	P30292 aspergillus
5	45	78.9	512	1 AMYA_DEBOC	P19269 debaryomyce
6	44	77.2	484	1 AMYA_ASPNG	P56271 aspergillus
7	39	68.4	975	1 KIT_CANPA	Q97799 canis fam1
8	39	68.4	978	1 KIT_FELCA	Q28889 felis silve
9	38	66.7	513	1 AMY2_SCHPO	Q14154 schizosach
10	37	64.9	494	1 AMY1_SACFI	P21567 saccharomyc
11	36	63.2	521	1 DRTS_MAIZE	Q81395 zea mays (m
12	36	63.2	1367	1 DPO3_THEMA	Q928f6 thermotoga
13	35	61.4	215	1 IF4E_APLCA	Q77710 aplysia cal
14	35	61.4	409	1 ARCA_BORAF	Q51896 borrelia af
15	35	61.4	410	1 ARCA_BORAF	Q51781 borrelia bu
16	35	61.4	410	1 HCTB_HELPPO	P12031 helix pomat
17	34	59.6	1390	1 N155_RAT	P37199 rattus norv
18	34	59.6	1391	1 N155_HUMAN	Q75694 homo sapien
19	33	57.9	200	1 YQAB_HAEIN	Q58366 haemophilus
20	33	57.9	260	1 YQ56_METJA	Q58366 methanococc
21	33	57.9	277	1 SP0A_CLOIN	P52839 clostridium
22	33	57.9	334	1 CUB_CLOPM	P29716 clostridium
23	33	57.9	352	1 FBRL_CABEL	Q22053 caenorhabdi
24	33	57.9	356	1 GATA_TREPA	Q83983 treponema p
25	33	57.9	525	1 VE59_LAMBD	P03754 bacterioph
26	33	57.9	530	1 DRTS_SOYBN	P51820 glycine max
27	33	57.9	539	1 YQ33_METJA	Q60356 methanococc
28	33	57.9	561	1 ILVD_METJA	Q58672 methanococc
29	33	57.9	607	1 YUKO_YEAST	P42951 saccharomyc
30	33	57.9	685	1 YG04_YEAST	P53118 saccharomyc
31	33	57.9	707	1 ETF2_MCV1	Q98277 molluscum c
32	33	57.9	710	1 ETF2_VACCA	Q57220 vaccinia vi
33	33	57.9	710	1 ETF2_VACCC	P20635 vaccinia vi

ALIGNMENTS

RESULT 1	AMYA_ASPAW	STANDARD;	PRT;	498 AA.
AC	002905;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Alpha-amylase A precursor (EC 3.2.1.1) (1.4-alpha-D-glucan			
DE	glucanohydrolase A).			
GN	AMYA.			
OS	Aspergillus awamori.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
OX	NCBI_TaxID=105351;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRATN-DVK143F;			
RK	MEDLINE=90254827; PubMed=2340591;			
RA	Korman D.R., Bayliss F.T., Barnett C.C., Cartmona C.L.,			
RA	Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,			
RA	Berka R.M.;			
RT	"Cloning, characterization, and expression of two alpha-amylase genes			
RT	from Aspergillus niger var. awamori."			
RU	Curr. Genet. 17:203-212(1990).			
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic			
CC	linkages in oligosaccharides and polysaccharides.			
CC	-1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO			
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-slb.ch/announce/			
CC	or send an email to license@isb-slb.ch).			
CC	-----			
DR	EMBL; X52755; CA36966.1; ..			
DR	HSSP; P10529; TTA.			
DR	InterPro; IPR000461; Alpha_amylase.			
DR	Pfam; PF00128; alpha-amylase; 1.			
KW	Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;			
KW	Glycoprotein; Signal; Multigene family.			
FT	SIGNAL 1 21			POTENTIAL.
FT	CHAIN 22 498			ALPHA-AMYLASE A.
FT	ACT_SITE 227 227			BY SIMILARITY.
FT	ACT_SITE 231 231			BY SIMILARITY.
FT	ACT_SITE 251 251			BY SIMILARITY.
FT	ACT_SITE 318 318			BY SIMILARITY.
FT	DISULFID 51 59			BY SIMILARITY.
FT	DISULFID 171 185			BY SIMILARITY.
FT	DISULFID 261 304			BY SIMILARITY.
FT	DISULFID 461 496			BY SIMILARITY.
FT	CARBOHYD 218 218			N-LINKED (GLCNAC. .) (POTENTIAL).
SEQUENCE	498 AA; 54880 MW; 7658511BC01A8A01 CRC64;			

34	33	57.9	710	1	ETP2_VACCT	O91f93 vaccinia vi
35	33	57.9	710	1	ETP2_VACCV	P20636 vaccinia vi
36	33	57.9	710	1	ETP2_VARY	P33806 variola vir
37	33	57.9	711	1	ETP2_MXXVL	O968k4 myxoma viru
38	33	57.9	711	1	ETP2_SPVKA	O968k4 myxoma viru
39	33	57.9	713	1	ETP2_YARBM	O968k4 shope fibro
40	33	57.9	775	1	POP1_SCHPO	O968k4 yaba monkey
41	33	57.9	1195	1	HEUS_METJA	P87060 schizosach
42	33	57.9	1458	1	CUT4_SCHPO	O58524 methanococc
43	33	57.9	1532	1	IGA_NEIGO	O9urv2 schizosach
44	33	57.9	1541	1	IGA1_HAEIN	P09790 neisseria g
45	33	57.9	1545	1	IGA3_HAEIN	P42782 haemophilus
						P45385 haemophilus

Query Match 100.0%; Score 57; DB 1; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.0011;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVDSYVFKEP 11
 |||||
 DB 151 SSVDSYVFKEP 161

RESULT 2

AMYA_ASPOP STANDARD; PRT; 499 AA.
 ID AMYA_ASPOP STANDARD; PRT; 499 AA.
 AC P10529; P11763; Q00250;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
 DE (1,4-alpha-D-glucan glucanohydrolase).
 GN AMY1 AND AMY2 AND AMY3.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_Taxid=5062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 63303;
 RX MEDLINE=89237897; PubMed=2785629;
 RA Wiersel S., Lachmund A., Wildhardt G., Rutkowski E.;
 RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
 RT intron-exon organization";
 RL Mol. Microbiol. 3:3-14(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89378767; PubMed=2789162;
 RA Genes M.J., Dove M.J., Selly V.L.;
 RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
 RT containing eight introns";
 RL Gene 79:107-117(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
 RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
 RT of Aspergillus oryzae";
 RL Agric. Biol. Chem. 53:593-599(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90128276; PubMed=2612911;
 RA Tsukagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A.,
 RA Ueki S.;
 RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
 RT evidence for multiple related genes";
 RL Gene 84:319-327(1989).
 RN [5]
 RP SEQUENCE OF 22-499.
 RA Toda H., Kondo K., Narita K.;
 RT "The complete amino acid sequence of Taka-amylase A";
 RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
 RN [6]
 RP SEQUENCE OF 206-225.
 RX MEDLINE=74001521; PubMed=4733850;
 RA Isemura S., Ikenaka T.;
 RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
 RT A with trypsin and chymotrypsin";
 RL J. Biochem. 74:1-10(1973).
 RN [7]
 RP SEQUENCE OF 433-499.
 RA Narita K.;
 RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 51:285-290(1975).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=80227691; PubMed=6156152;

RA Matsura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
 RA Toda H., Narita K., Kakudo M.;
 RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
 RT 3-A resolution";
 RL J. Biochem. 87:1555-1558(1980).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=84212370; PubMed=6609921;
 RA Matsura Y., Kusunoki M., Harada W., Kakudo M.;
 RT "Structure and possible catalytic residues of Taka-amylase A";
 RL J. Biochem. 95:697-702(1984).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
 RX MEDLINE=97428212; PubMed=9283074;
 RA Brozowski A.M., Davies G.J.;
 RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
 RT inhibitor acarbose at 2.0-A resolution";
 RL Biochemistry 36:10837-10845(1997).
 CC -I- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -I- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -I- MISCELLANEOUS: THE SEQUENCE OF AMY1 AND AMY2 IS SHOWN.
 CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X12725; CAA31218.1; -;
 DR EMBL: X12726; CAA31219.1; -;
 DR EMBL: X12727; CAA31220.1; -;
 DR EMBL: D00434; BAA00336.1; -;
 DR EMBL: M33218; AAA32708.1; -;
 DR PIR: S04548; ALAS1.
 DR PIR: S04549; ALAS3.
 DR PIR: A10627; ALAS3.
 DR PIR: JK0201; JK0201.
 DR PIR: JS0240; JS0240.
 DR PIR: JT0466; JT0466.
 DR PDB: 2TAA; 15-OCT-89.
 DR PDB: 6TAA; 31-OCT-93.
 DR PDB: 7TAA; 25-NOV-98.
 DR Glycosultdb: P10529; -;
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR HydroLase: Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family; 3D-structure.
 FT STGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 FT
 FT VARIANT 56 56
 FT VARIANT 172 172
 FT VARIANT 93 94
 FT CONFLICT 106 106
 FT CONFLICT 184 184
 FT CONFLICT 195 195
 FT CONFLICT 255 255
 FT CONFLICT 291 291
 FT CONFLICT 345 345
 N-LINKED (GLCNAC. . .).
 /FTID=CAR_000123.
 O -> R (IN AMY3).
 F -> L (IN AMY3).
 TT -> DC (IN REF. 5).
 O -> T (IN REF. 5).
 D -> Y (IN REF. 3).
 P -> L (IN REF. 3).
 D -> V (IN REF. 3).
 D -> H (IN REF. 4).
 I -> L (IN REF. 5).

```

FT CONFLICT 370 370 L -> A (IN REF. 4).
FT CONFLICT 406 409 WPXY -> PYI (IN REF. 5).
FT CONFLICT 448 448 G -> S (IN REF. 5).
FT CONFLICT 497 497 S -> SD (IN REF. 5 AND 7).
FT HELIX 24 27
FT TURN 28 29
FT STRAND 32 35
FT HELIX 37 40
FT STRAND 42 42
FT TURN 43 44
FT HELIX 53 55
FT STRAND 61 61
FT HELIX 63 67
FT TURN 68 69
FT HELIX 70 74
FT TURN 75 77
FT STRAND 80 83
FT STRAND 87 89
FT STRAND 94 94
FT TURN 95 96
FT STRAND 97 97
FT TURN 100 101
FT STRAND 105 111
FT TURN 113 115
FT TURN 118 119
FT HELIX 120 130
FT TURN 131 132
FT STRAND 134 139
FT STRAND 143 143
FT STRAND 146 146
FT HELIX 150 152
FT TURN 155 156
FT STRAND 159 159
FT HELIX 164 166
FT STRAND 167 167
FT STRAND 172 172
FT HELIX 179 184
FT STRAND 186 188
FT STRAND 193 194
FT STRAND 196 197
FT TURN 199 200
FT HELIX 202 219
FT TURN 220 220
FT STRAND 223 226
FT TURN 227 228
FT HELIX 229 231
FT TURN 234 235
FT HELIX 236 244
FT STRAND 247 250
FT HELIX 257 265
FT TURN 266 266
FT STRAND 269 271
FT HELIX 273 283
FT TURN 286 287
FT HELIX 290 303
FT STRAND 307 309
FT HELIX 311 312
FT TURN 317 318
FT HELIX 322 325
FT STRAND 329 341
FT TURN 345 349
FT STRAND 350 351
FT TURN 352 354
FT HELIX 355 355
FT TURN 361 364
FT HELIX 368 371
FT TURN 372 372

```

Query Match 100.0%; Score 57; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.0012;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11

```

Db 151 SSVDSVFKPF 161
|||||
RESULT 3
AMYB_ASPAW STANDARD; PRT; 499 AA.
AC 002906;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alpha-amylase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase B).
GN AMYB.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxId=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UVK143F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berka R.M.;
RT *Cloning, characterization, and expression of two alpha-amylase genes
RT from Aspergillus niger var. awamori.;
RL Curr. Genet. 17:203-212(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X52756; CAA36967.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR00461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 499
FT ACT_SITE 227 227
FT ACT_SITE 231 231
FT ACT_SITE 251 251
FT ACT_SITE 318 318
FT DISULFID 51 59
FT DISULFID 171 185
FT DISULFID 261 304
FT DISULFID 461 496
FT CARBOHYD 218 218
SQ SEQUENCE 499 AA; 54921 MW; 740B96B1BC01A8A CRC64;

```

Query Match 100.0%; Score 57; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.0012;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11

Db 151 SSVDSVFKPF 161

RESULT 4
 AMY_ASPSH

```

ID  AMY_ASPSH          STANDARD:          PRT:          499 AA.
AC  P30292;
DT  01-APR-1993 (Rel. 25, Created)
DT  01-APR-1993 (Rel. 25, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE  glucanohydrolase).
GN  AMY.
OS  Aspergillus shirousamli.
OC  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC  Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX  NCBI_TaxID=5070;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92323146; PubMed=1368777;
RA  Shibuya I., Tamura G., Ishikawa T., Hara S.;
RT  "Cloning of the alpha-amylase cDNA of Aspergillus shirousamli and its
RT  expression in Saccharomyces cerevisiae."
RL  Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC  -I- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC  linkages in oligosaccharides and polysaccharides.
CC  -I- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC  -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; D10461; BAA01255.1; -.
DR  PIR; J50663; JS0663.
DR  HSSP; P10529; 7TAA.
DR  InterPro: IPR000461; Alpha-amylase.
DR  Pfam: PF00128; alpha-amylase; 1.
KW  Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW  Glycoprotein; Signal.
FT  CHAIN 1
FT  SIGNAL 21
FT  ACT_SITE 22 499 ALPHA-AMYLASE.
FT  ACT_SITE 227 227
FT  ACT_SITE 231 231 BY SIMILARITY.
FT  ACT_SITE 251 251 BY SIMILARITY.
FT  ACT_SITE 318 318 BY SIMILARITY.
FT  DISULFID 51 59 BY SIMILARITY.
FT  DISULFID 171 185 BY SIMILARITY.
FT  DISULFID 261 304 BY SIMILARITY.
FT  DISULFID 461 496 BY SIMILARITY.
FT  CARBOHYD 218 218 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ  SEQUENCE 499 AA; 54852 MW; 1FB7AE5DDA01C03F CRC64;

```

```

Query Match          100.0%; Score 57; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY  1 SSVDSYVKPF 11
    |||||
DB  151 SSVDSYVKPF 161

```

```

RESULT 5
AMY1_DEBOC          STANDARD:          PRT:          512 AA.
AC  P19269;
DT  01-NOV-1990 (Rel. 16, Created)
DT  01-NOV-1990 (Rel. 16, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Alpha-amylase 1 precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE  glucanohydrolase).
GN  AMY1.

```

```

OS  Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC  Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX  NCBI_TaxID=27300;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=90032659; PubMed=2806251;
RC  STRAIN=ATCC 26076;
RA  Strasser A.W.M., Selk R., Dohmen R.J., Niemann T., Bielefeld M.,
RT  "Analysis of the alpha-amylase gene of Schwanniomyces occidentalis
RT  and the secretion of its gene product in transformants of different
RT  yeast genera."
RL  Eur. J. Biochem. 184:699-706(1989).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  STRAIN=CCRC 21164;
RC  MEDLINE=92120467; PubMed=1769525;
RA  Wu F.M., Wang T.T., Hsu W.H.;
RT  "The nucleotide sequence of Schwanniomyces occidentalis alpha-amylase
RT  gene."
RL  FEMS Microbiol. Lett. 66:313-318(1991).
RN  [3]
RP  SEQUENCE FROM N.A.
RX  STRAIN=ATCC 26077 / CBS 2863;
RC  MEDLINE=92307400; PubMed=1612414;
RA  Park J.C., Bai S., Tai C.Y., Chun S.B.;
RT  "Nucleotide sequence of the extracellular alpha-amylase gene in the
RT  yeast Schwanniomyces occidentalis ATCC 26077."
RL  FEMS Microbiol. Lett. 72:17-23(1992).
CC  -I- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC  linkages in oligosaccharides and polysaccharides.
CC  -I- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC  -I- ENZYME REGULATION: ALPHA-AMYLASE EXPRESSION UNDERLIES CATABOLITE
CC  REPRESSION BY GLUCOSE.
CC  -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; S77586; AAB21151.2; -.
DR  EMBL; X16040; CAA34162.1; -.
DR  EMBL; X62079; CAA43995.1; -.
DR  EMBL; S38381; AAB22383.2; -.
DR  PIR; S06115; S06115.
DR  PIR; S23355; S23355.
DR  HSSP; P10529; 7TAA.
DR  InterPro: IPR000461; Alpha-amylase.
DR  Pfam: PF00128; alpha-amylase; 1.
KW  Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW  Glycoprotein; Signal.
FT  CHAIN 1
FT  SIGNAL 25
FT  ACT_SITE 26 512 POTENTIAL.
FT  ACT_SITE 242 242 ALPHA-AMYLASE 1.
FT  ACT_SITE 246 246 BY SIMILARITY.
FT  ACT_SITE 333 333 BY SIMILARITY.
FT  CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (PROBABLE).
FT  DISULFID 66 74 BY SIMILARITY.
FT  DISULFID 186 200 BY SIMILARITY.
FT  DISULFID 276 319 BY SIMILARITY.
FT  DISULFID 475 510 BY SIMILARITY.
FT  VARIANT 32 32 M -> K (IN STRAINS CCRC 21164 AND ATCC
FT  VARIANT 36 36 S -> G (IN STRAIN CCRC 21164).
FT  VARIANT 73 73 Y -> I (IN STRAIN ATCC 26077).
FT  VARIANT 280 280 Y -> S (IN STRAIN CCRC 21164).
FT  VARIANT 350 350 D -> A (IN STRAINS CCRC 21164 AND ATCC

```



```

FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 975 AA; 109335 MW; 8F570BD9F05B1CB CRC64;

Query Match 68.4%; Score 39; DB 1; Length 975;
Best Local Similarity 77.8%; Pred. No. 8.4;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSVDSYVEK 9
Db 479 SSIDYSAFK 487

RESULT 8
KIT_FELCA STANDARD; PRT; 978 AA.
AC 028889;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Mest/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT.
OS Felis silvestris catus (cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxId=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RA MEDLINE=95140426; PubMed=7530827;
RA Herbst R., Mumentstu S., Ullrich A.;
RT "oncogenic activation of v-kit involves deletion of a putative
RT tyrosine-substrate interaction site.";
RL Oncogene 10:369-379(1995).
CC -1- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S76596; AAB33207.1; -.
DR HSSP; P11362; IAGW.
DR InterPro; IPR000719; Euk-Pkinase.
DR InterPro; IPR003599; I9.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003600; IG_1ike.
DR InterPro; IPR001824; Receptor tyr_kin_III.
DR InterPro; IPR001245; Tyr-kinase.
DR Pfam; PF00047; Ig; 2.
DR Pfam; PF00069; pkinase; 2.
DR SMART; SM00409; IG; 2.
DR SMART; SM00410; IG_1ike; 1.
DR SMART; SM00219; TyrcK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

```

```

DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 978 MAST/STEM CELL GROWTH FACTOR RECEPTOR.
FT DOMAIN 23 521 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 522 544 POTENTIAL.
FT DOMAIN 545 978 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 590 938 PROTEIN KINASE.
FT NP_BIND 596 604 ATP (BY SIMILARITY).
FT BINDING 624 624 ATP (BY SIMILARITY).
FT ACT_SITE 793 793 BY SIMILARITY.
FT MOD_RES 824 824 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 978 AA; 109449 MW; 6D45472E07440B6B CRC64;

Query Match 68.4%; Score 39; DB 1; Length 978;
Best Local Similarity 77.8%; Pred. No. 8.4;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSVDSYVEK 9
Db 477 SSIDYSAFK 485

RESULT 9
AMY2_SCHPO STANDARD; PRT; 513 AA.
ID AMY2_SCHPO
AC 014154;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Probable alpha-amylase C4A8.01 precursor (EC 3.2.1.1) (1,4-alpha-D-
DE glucan glucanhydrolase).
GN SPAC4A8.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxId=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RA Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
RL -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
RL linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: Z98762; CAB11471.1; -
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
 FT SIGNAL 1 25
 FT CHAIN 26 513
 FT ACT_SITE 226 226 PROBABLE ALPHA-AMYLASE C4A8.01.
 FT ACT_SITE 230 230 BY SIMILARITY.
 FT ACT_SITE 250 250 BY SIMILARITY.
 FT ACT_SITE 318 318 BY SIMILARITY.
 FT DISULFID 52 60 BY SIMILARITY.
 FT DISULFID 171 184 BY SIMILARITY.
 FT DISULFID 260 304 BY SIMILARITY.
 FT DISULFID 454 488 BY SIMILARITY.
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 513 AA: 58715 MW: 455097F8428C182 CRC64;

Query Match 66.7%; Score 38; DB 1; Length 513;
 Best Local Similarity 40.0%; Pred. No. 6.7;
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 SVDSYVFKPF 11
 ::||::|||

DB 152 NIDYGIYRPF 161

RESULT 10
 AMY1_SACFI STANDARD: PRT; 494 AA.
 ID AMY1_SACFI P21567;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Alpha-amylase precursor (EC 3.2.1.1).
 GN ALPL.
 OS Saccharomycopsis fibuligera (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
 OX NCBI_TaxID=4944;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87276512; PubMed=3497057;
 RA Itoh T., Yamashita I., Fukui S.;
 RT "Nucleotide sequence of the alpha-amylase gene (ALPL) in the yeast
 Saccharomycopsis fibuligera.";
 RL FEBS Lett. 219:339-342(1987).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES. ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X05791; CAA29233.1; -
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02806; alpha-amylase_C; 1.
 DR Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal.
 FT SIGNAL 1 26 POTENTIAL.

FT CHAIN 27 494
 FT ACT_SITE 233 233 ALPHA-AMYLASE.
 FT ACT_SITE 237 237 BY SIMILARITY.
 FT ACT_SITE 324 324 BY SIMILARITY.
 FT CARBOHYD 224 224 BY SIMILARITY.
 FT DISULFID 57 65 N-LINKED (GLCNAC. . .) (PROBABLE).
 FT DISULFID 177 191 BY SIMILARITY.
 FT DISULFID 267 310 BY SIMILARITY.
 FT DISULFID 462 493 BY SIMILARITY.
 SQ SEQUENCE 494 AA: 54387 MW: 77788FDD6BF58B67 CRC64;

Query Match 64.9%; Score 37; DB 1; Length 494;
 Best Local Similarity 60.0%; Pred. No. 10;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SVDSYVFKPF 11
 ::|||::||

DB 158 SIDYSEYRPF 167

RESULT 11
 DRTS_MAIZE STANDARD: PRT; 521 AA.
 ID DRTS_MAIZE O81395;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Bifunctional dihydrofolate reductase-thymidylate synthase (DHFR-
 DE [includes: Dihydrofolate reductase (EC 1.5.1.3); Thymidylate synthase
 DE (EC 2.1.1.45)].
 GN DRTS.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cox R.M., Robertson D., Fites R.C.;
 RT "Cloning of a cDNA encoding the bifunctional dihydrofolate reductase-
 RT thymidylate synthase (DHFR-*ts*) in Zea mays.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: 5,6,7,8-tetrahydrofolate + NADP(+) = 7,8-
 CC dihydrofolate + NADPH.
 CC -1- CATALYTIC ACTIVITY: 5,10-methylenetetrahydrofolate + dUMP =
 CC dihydrofolate + dTMP.
 CC -1- PATHWAY: ESSENTIAL STEP FOR DE NOVO GLYCINE AND PURINE SYNTHESIS,
 CC DNA PRECURSOR SYNTHESIS, AND FOR THE CONVERSION OF DUMP TO dTMP.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE
 CC DIHYDROFOLATE REDUCTASE FAMILY.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE THYMIDYLATE
 CC SYNTHASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF073488; AAC26003.1; -
 DR HSSP: P13100; 1C17.
 DR Mendel: 32032; Zeama: 2610; 32032.
 DR InterPro: IPR001796; DHFR.
 DR InterPro: IPR000398; Thymidylat_synth.
 DR Pfam: PF00186; Dihfolate_red; 1.
 DR Pfam: PF00303; thymidylat_synth; 1.
 DR PRINTS: PR00070; DHFR.
 DR PRINTS: PR00108; THYMSNTHASE.
 DR Prodom: PD001180; Thymidylat_synth; 1.
 DR PROSITE: PS00075; DHFR; 1.

DR PROSITE: PS00091: THYMIDYLATE SYNTHASE; 1.
 KW Multifunctional enzyme; Oxidoreductase; Transferase; NADP:
 KM Methyltransferase; Nucleotide biosynthesis; One-carbon metabolism.
 FT DOMAIN 1 196 DIHYDROFOLATE REDUCTASE.
 FT DOMAIN 197 521 THYMIDYLATE SYNTHASE.
 FT ACT_SITE 403 403 BY SIMILARITY.
 SO SEQUENCE 521 AA; 58965 MW; 81266f8652625f06 CRC64;

Query Match 63.2%; Score 36; DB 1; Length 521;
 Best Local Similarity 66.7%; Pred. No. 17;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VDYSVFKPF 11
 11:111:11
 Db 166 VDYSVFKPF 174

RESULT 12
 ID DPO3_THEMA STANDARD; PRT; 1367 AA.
 AC Q9ZHF6;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA polymerase III polC-type (EC 2.7.7.7) (PolIII).
 GN POLC OR TM0576.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99045593; PubMed=9826752;
 RA Huang Y.P., Ito J.;
 RT "The hyperthermophilic bacterium Thermotoga maritima has two different
 RT classes of family C DNA polymerases: evolutionary implications.";
 RT Nucleic Acids Res. 26:5300-5309(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Hat D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Uitterlind T.R., Malek J.A., Linher C.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RA Salzberg S.L.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima.";
 RT Nature 399:323-329(1999).
 RL NATURE 399:323-329(1999).
 CC -1- FUNCTION: REQUIRED FOR REPLICATIVE DNA SYNTHESIS. THIS DNA
 CC POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
 CC + [DNA](N).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-C FAMILY. POLC
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF065313; AAC80438.1; -;
 CC EMBL: AE001732; AAD35661.1; -;
 CC TIGR: TM0576; -;
 DR InterPro: IPR002562; 3_5_exonuclease.
 DR InterPro: IPR004013; PHP_C.
 DR InterPro: IPR003141; PHP_N.

DR InterPro: IPR002309; tRNA-synL-2.
 DR Pfam: PF00929; Exonuclease; 1.
 DR Pfam: PF02811; PHP_C; 1.
 DR Pfam: PF02231; PHP_N; 1.
 DR Pfam: PF01336; tRNA_ant1; 1.
 DR SMART: SM00474; 35EXOC; 1.
 DR SMART: SM00479; EXOIII; 1.
 DR SMART: SM00481; POLIITAC; 1.
 KW Transferase; DNA-directed DNA polymerase; DNA replication; Hydrolase;
 FT DOMAIN 358 521 EXONUCLEASE.
 SO SEQUENCE 1367 AA; 155362 MW; EE5916FA0591F84 CRC64;

Query Match 63.2%; Score 36; DB 1; Length 1367;
 Best Local Similarity 50.0%; Pred. No. 47;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 SVDSVFKPF 11
 ::11:111
 Db 537 TIDYALKEP 546

RESULT 13
 ID IF4E_APLCA STANDARD; PRT; 215 AA.
 AC 077210;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Eukaryotic translation Initiation factor 4E (eIF4E) (mRNA
 DE cap-binding protein) (eIF-4F 25 kDa subunit).
 OS Aplysia californica (California sea hare).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia; Anasplidae;
 OX Aplysiidae; Aplysia.
 RN NCBI_TaxID=6500;
 RN [1]
 RP SEQUENCE FROM N.A., AND PHOSPHORYLATION.
 RC TISSUE=Nerve;
 RX MEDLINE=99009053; PubMed=9792652;
 RA Dyer J.R., Peipo A.M., Yanow S.K., Sossin W.S.;
 RT "Phosphorylation of eIF4E at a conserved serine in Aplysia.";
 RT J. Biol. Chem. 273:29469-29474(1998).
 RL J. Biol. Chem. 273:29469-29474(1998).
 CC -1- FUNCTION: RECOGNIZES AND BINDS THE 7-METHYLGUANOSINE-CONTAINING
 CC MRNA "CAP" DURING AN EARLY STEP IN THE INITIATION OF PROTEIN
 CC SYNTHESIS AND FACILITATES RIBOSOME BINDING BY INDUCING THE
 CC UNWINDING OF THE MRNA SECONDARY STRUCTURES.
 CC -1- SUBUNIT: EIF4E IS A TRIMER COMPOSED OF EIF4E, EIF4G AND EIF4A
 CC (WHICH CAN CYCLE IN AND OUT OF THE COMPLEX).
 CC -1- PTM: PHOSPHORYLATION INCREASES THE ABILITY OF THE PROTEIN TO BIND
 CC TO MRNA CAPS AND TO FORM THE EIF4F COMPLEX.
 CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF085810; AAC36720.1; -;
 CC HSSP: P07260; IAP8.
 DR InterPro: IPR001040; eIF_4E.
 DR Pfam: PF01652; IF4E; 1.
 DR PRODOM: PD003697; eIF_4E; 1.
 DR PROSITE: PS00813; IF4E; 1.
 KW Initiation factor; Protein biosynthesis; RNA-binding; Phosphorylation.
 FT MOD_RES 207 207 PHOSPHORYLATION (BY PKC).
 SO SEQUENCE 215 AA; 24646 MW; 9EE00CB6DE812E7 CRC64;

Query Match 61.4%; Score 35; DB 1; Length 215;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:55:49 ; Search time 73.31 Seconds
(without alignments)

14.418 Million cell updates/sec

Title: US-09-710-339-2_COPY_150_160

Perfect score: 57

Sequence: 1 SSVDSYVFKPF 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

PIR_71:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	478	2 JK0201	alpha-amylase (EC
2	57	100.0	498	2 A48305	alpha-amylase (EC
3	57	100.0	499	1 ALA51	alpha-amylase (EC
4	57	100.0	499	1 ALA51	alpha-amylase (EC
5	57	100.0	499	2 B48305	alpha-amylase (EC
6	57	100.0	499	2 JS0663	alpha-amylase (EC
7	57	100.0	499	2 JT0466	alpha-amylase (EC
8	57	100.0	499	2 JS0588	alpha-amylase (EC
9	45	78.9	512	2 S06115	alpha-amylase (EC
10	45	78.9	512	2 S23355	alpha-amylase (EC
11	44	77.2	484	1 A35282	alpha-amylase (EC
12	40	70.2	489	2 AF3457	D-hydantoinase (EC
13	39	68.4	4116	2 T13719	calo protein - fru
14	38	66.7	425	2 H82183	conserved hypotet
15	38	66.7	430	2 T37549	hypothetical prote
16	38	66.7	513	2 T38770	alpha-amylase a pr
17	38	66.7	1869	2 A59290	class V chitin syn
18	37	64.9	494	1 ALBVAF	alpha-amylase (EC
19	36	63.2	112	2 T47522	hypothetical prote
20	36	63.2	173	2 T27650	hypothetical prote
21	36	63.2	300	2 G86765	hypothetical prote
22	36	63.2	353	2 T19522	hypothetical prote
23	36	63.2	409	2 T46063	hypothetical prote
24	36	63.2	491	2 T38448	probable alpha-am
25	36	63.2	521	2 T01684	dihydrofolate redu
26	36	63.2	541	2 T43862	FAD flavoprotein o
27	36	63.2	624	2 T39102	hypothetical prote
28	36	63.2	1367	2 C72360	DNA polymerase III
29	36	63.2	2692	2 T23768	hypothetical prote

30	35	61.4	116	2 H98221	hypothetical prote
31	35	61.4	153	2 AH3064	hypothetical prote
32	35	61.4	363	2 T44564	probable thiamin b
33	35	61.4	377	2 T50563	beta-1,3-glucanase
34	35	61.4	410	2 A29393	hemocyanin beta-c
35	35	61.4	410	2 H70204	arginine deiminase
36	35	61.4	492	2 F86384	probable protein f
37	35	61.4	497	2 T52308	very-long-chain fa
38	35	61.4	562	2 T34319	hypothetical prote
39	35	61.4	1498	2 S78102	chitin synthase (E
40	35	61.4	1852	2 JC5546	chitin synthase (E
41	34	59.6	201	2 AH2231	hypothetical prote
42	34	59.6	213	2 H83839	hypothetical prote
43	34	59.6	370	2 C71052	hypothetical prote
44	34	59.6	407	2 T30469	hypothetical prote
45	34	59.6	418	2 E96687	hypothetical prote

ALIGNMENTS

RESULT 1

JK0201
alpha-amylase (EC 3.2.1.1) - Aspergillus oryzae

N:Alternate names: glycogenase; Taka-amylase A

C:Species: Aspergillus oryzae

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998

C:Accession: JK0201

R:Yoda, H.; Kondo, K.; Narita, K.
Proc. Jpn. Acad. 58B, 208-212, 1982

A:Title: The complete amino acid sequence of Taka-amylase A.

A:Reference number: JK0201

A:Molecule type: protein

A:Residues: 1-478 <100>

C:Comment: One atom of calcium per molecule is essential for the activity.

C:Comment: This enzyme is a glycoprotein.

C:Comment: See also PIR:JT0466 and PIR:JS0240.

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:173-300/Domains: alpha-amylase core homology <AMY>

F:197/Binding site: carboxylate (Asn) (covalent) #status experimental

F:210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 57; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0049;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSVDSYVFKPF 11
|||||

DB 130 SSVDSYVFKPF 140

RESULT 2
alpha-amylase (EC 3.2.1.1) A precursor - Aspergillus awamori

C:Species: Aspergillus awamori

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997

C:Accession: A48305

R:Kotman, D.R.; Bayliss, F.T.; Barnett, C.C.; Camona, C.L.; Kodana, K.H.; Royer, T.J.
Curr. Genet. 17, 203-212, 1990

A:Title: Cloning, characterization, and expression of two alpha-amylase genes from As

A:Reference number: A48305; MUID:90254827

A:Accession: A48305

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-498 <KOR>

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A.Pathway: glycogen/starch degradation
 C.Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C.Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F.194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 57; DB 2; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 SSVDSVFKPF 11
 |||||||||
 Db 151 SSVDSVFKPF 161

RESULT 3
 ALAS1
 alpha-amylase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae
 N.Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A
 C.Species: Aspergillus oryzae
 C.Date: 30-Sep-1991 #sequence-revision 30-Sep-1991 #text-change 15-Sep-2000
 C.Accession: S04548; A33214; J50240; A91930; A93767; A10627
 R.Witzel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989
 A.Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon
 A.Reference number: S04548; MUID:89237897
 A.Accession: S04548
 A.Molecule type: DNA
 A.Residues: 1-499 <WIR>
 A.Cross-references: EMBL:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921
 A.Genetics: AMY1
 A.Accession: A33214
 A.Molecule type: mRNA
 A.Residues: 1-499 <MID>
 A.Cross-references: GB:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921
 R.Genes, M.U.; Dove, M.U.; Seligy, V.L.
 Gene 79, 107-117, 1989
 A.Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
 A.Reference number: J50240; MUID:89378767
 A.Accession: J50240
 A.Molecule type: DNA
 A.Residues: 1-499 <GEN>
 A.Genetics: AMY2
 A.Note: the authors refer to this as isozyme II
 R.Isemura, S.; Ikenaka, T.
 J. Biochem. 74, 1-10, 1973
 A.Reference number: A91930; MUID:74001521
 A.Accession: A91930
 A.Molecule type: protein
 A.Residues: 206-225 <ISE>
 R.Narita, K.
 Proc. Jpn. Acad. 51, 285-290, 1975
 A.Reference number: A93767
 A.Accession: A93767
 A.Molecule type: protein
 A.Residues: 434-443, 446-447, 'Q', 449-458, 'GRV', 459-464, 467-468, 'B', 470, 'B', 472-499 <NMR>
 J. Biochem. 95, 697-702, 1984
 A.Title: Structure and possible catalytic residues of Taka-amylase A.
 A.Reference number: A37454; MUID:8421370
 A.Contents: annotation: X-ray crystallography, 3.0 angstroms
 R.Swift, H.J.; Brady, L.; Dewende, Z.S.; Dodson, E.J.; Turkenburg, J.P.; Wilkinson, A.
 submitted to the Brookhaven Protein Data Bank, August 1992
 A.Reference number: A51548; PDB:6RTA
 A.Contents: annotation: X-ray crystallography, 2.1 angstroms, residues 22-497
 C.Comment: One atom of calcium per molecule is essential for activity.
 C.Genetics: <AMY1>
 A.Gene: amy1
 A.Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C.Genetics: <AMY2>
 A.Gene: amy2; AmyII
 A.Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C.Function:

A.Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A.Pathway: glycogen/starch degradation
 C.Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C.Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; po
 F.1-21/Domain: signal sequence #status predicted <SIG>
 F.22-499/Product: alpha-amylase 1 #status experimental <MAT>
 F.194-321/Domain: alpha-amylase core homology <AMY>
 F.51-59, 171-185, 261-304, 461-496/Disulfide bonds: #status experimental
 F.142, 183, 196, 231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F.218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F.227, 251, 318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 57; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 SSVDSVFKPF 11
 |||||||||
 Db 151 SSVDSVFKPF 161

RESULT 4
 ALAS3
 alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae
 N.Alternate names: alpha-amylase isozyme I; glycogenase; Taka-amylase A
 C.Species: Aspergillus oryzae
 C.Date: 30-Sep-1991 #sequence-revision 30-Sep-1991 #text-change 18-Jun-1999
 C.Accession: S04549; A33215; A44713
 R.Witzel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989
 A.Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exo
 A.Reference number: S04549; MUID:89237897
 A.Accession: S04549
 A.Molecule type: DNA
 A.Residues: 1-499 <WIR>
 A.Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 A.Accession: A33215
 A.Status: translation not shown
 A.Molecule type: mRNA
 A.Residues: 1-499 <MID>
 A.Cross-references: GB:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 R.Genes, M.U.; Dove, M.U.; Seligy, V.L.
 Gene 79, 107-117, 1989
 A.Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each contain
 A.Reference number: J50240; MUID:89378767
 A.Accession: A44713
 A.Molecule type: DNA
 A.Residues: 1-499 <GEN>
 A.Note: the authors refer to this as isozyme I
 R.Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
 J. Biochem. 95, 697-702, 1984
 A.Title: Structure and possible catalytic residues of Taka-amylase A.
 A.Reference number: A37454; MUID:8421370
 A.Contents: annotation: X-ray crystallography, 3.0 angstroms
 C.Comment: One atom of calcium per molecule is essential for activity.
 C.Genetics: <AMY1>
 A.Gene: amy1
 A.Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C.Function:
 A.Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A.Pathway: glycogen/starch degradation
 C.Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C.Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me
 F.1-21/Domain: signal sequence #status predicted <SIG>
 F.22-499/Product: alpha-amylase 3 #status experimental <MAT>
 F.194-321/Domain: alpha-amylase core homology <AMY>
 F.51-59, 171-185, 261-304, 461-496/Disulfide bonds: #status experimental
 F.142, 183, 196, 231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F.218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F.231, 251, 318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 57; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11
|||||
DB 151 SSVDSVFKPF 161

RESULT 5

B48305

alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori

C:Species: Aspergillus awamori
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997

C:Accession: B48305

R:Korman, D. R.; Bayliss, F. T.; Barnett, C. C.; Carmona, C. L.; Kodama, K. H.; Royer, T. J.;
Curr. Genet. 17, 203-212, 1990A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus
A:Reference number: M48305; MUID:90254827

A:Accession: B48305

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-499 <KOR>

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 57; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11
|||||
DB 151 SSVDSVFKPF 161

RESULT 6

J50663

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.

C:Species: Aspergillus sp.

C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999

C:Accession: J50663

R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
Biotech. Biotechnol. Biochem. 56, 174-179, 1992

A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shiroyamii and its expression

A:Reference number: J50663; MUID:92323146

A:Accession: J50663

A:Molecule type: mRNA

A:Residues: 1-499 <SHR>

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-499/Product: alpha-amylase #status predicted <AMP>

F:194-321/Domain: alpha-amylase core homology <AMY>

F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 57; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11
|||||
DB 151 SSVDSVFKPF 161

RESULT 7
JT0466
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae

N:Alternate names: glycosidase; Taka-amylase A

C:Species: Aspergillus oryzae

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998

C:Accession: JT0466

R:Tada, S.; Jimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
Agric. Biol. Chem. 53, 593-599, 1989

A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Aspergillus

A:Reference number: JT0466

A:Accession: JT0466

A:Molecule type: DNA

A:Residues: 1-499 <TAD>

C:Comment: See also PIR:JK0201 and PIR:J50240.

C:Comment: One atom of calcium per molecule is essential for activity.

C:Genetics:

A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-499/Product: alpha-amylase #status predicted <MAT>

F:194-321/Domain: alpha-amylase core homology <AMY>

F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 57; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11
|||||
DB 151 SSVDSVFKPF 161

RESULT 8

JN0588

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae

N:Alternate names: Taka-amylase A

C:Species: Aspergillus oryzae

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997

C:Accession: JN0588

R:Tukagoshi, N.; Furukawa, M.; Nagaba, H.; Kirita, N.; Tsuboi, A.; Uchida, S.
Gene 84, 319-327, 1989

A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for

A:Reference number: JN0588; MUID:90128276

A:Accession: JN0588

A:Molecule type: mRNA

A:Residues: 1-499 <TSU>

C:Comment: The alpha amylases are encoded by multigene family.

C:Genetics:

A:Gene: Taa-G1

A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-499/Product: alpha-amylase #status predicted <MAT>

F:194-321/Domain: alpha-amylase core homology <AMY>

F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 57; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFKPF 11

Db 151 SSVDSYVFKPF 161

RESULT 9

S06115
alpha-amylase (EC 3.2.1.1) AMY1 precursor - yeast (Schwanniomyces occidentalis)
C:Species: Schwanniomyces occidentalis
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 22-Jun-1999
C:Accession: S06115
R:Strasser, A.W.M.; Selk, R.; Dohmen, R.J.; Niemann, T.; Bielefeld, M.; Seeboth, F.; Tu
Eur. J. Biochem. 184, 699-706, 1999
A:Title: Analysis of the alpha-amylase gene of Schwanniomyces occidentalis and the secret
A:Reference number: S06115; MUID:90032659
A:Accession: S06115
A:Molecule type: DNA
A:Residues: 1-512 <STR>
A:Cross-references: EMBL:X16040; NID:g4882; PIDN:CNA34162.1; PID:g4883
C:Genetics:
A:Gene: AMY1
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-512/Product: alpha-amylase #status predicted <MAT>
F:209-336/Domain: alpha-amylase core homology <AMY>

Query Match 78.9%; Score 45; DB 2; Length 512;
Best Local Similarity 81.8%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSVDSYVFKPF 11
Db 166 SSVDSYVFKPF 176

RESULT 10
S23355
alpha-amylase (EC 3.2.1.1) precursor - yeast (Schwanniomyces occidentalis)
C:Species: Schwanniomyces occidentalis
C>Date: 22-Nov-1993 #sequence_revision 13-Mar-1997 #text_change 22-Jun-1999
C:Accession: S23355
R:Yu, F.M.; Wang, T.T.; Hsu, W.H.
FEMS Microbiol. Lett. 82, 313-318, 1991
A:Title: The nucleotide sequence of Schwanniomyces occidentalis alpha-amylase gene.
A:Reference number: S23355
A:Accession: S23355
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-512 <WUR>
A:Cross-references: EMBL:X62079; NID:g4880; PIDN:CAA43995.1; PID:g4881
A>Note: the authors translated the codon AGA for residue 21 as Pro, CTT for residue 61
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase
F:209-336/Domain: alpha-amylase core homology <AMY>

Query Match 78.9%; Score 45; DB 2; Length 512;
Best Local Similarity 81.8%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSVDSYVFKPF 11
Db 166 SSVDSYVFKPF 176

RESULT 11
A35282
alpha-amylase (EC 3.2.1.1) - Aspergillus niger
C:Species: Aspergillus niger

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A35282
R:Boel, E.; Brady, L.; Brzozowski, A.M.; Derewenda, Z.; Dodson, G.G.; Jensen, V.J.; P
Biochemistry 29, 6244-6249, 1990
A:Title: Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-angstrom
A:Reference number: A35282; MUID:91002514
A:Accession: A35282
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-484 <BOE>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:173-300/Domain: alpha-amylase core homology <AMY>

Query Match 77.2%; Score 44; DB 1; Length 484;
Best Local Similarity 72.7%; Pred. No. 1.6;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSVDSYVFKPF 11
Db 130 NDVDSYVFKPF 140

RESULT 12
AF3457
D-hydantoinase (EC 3.5.2.2) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AF3457
R:Delvecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujter, C.; Ios, T.; Ivanov
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3457
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-489 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52825.1; PID:g17983665; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11644
A:Map position: I
A:Superfamily: allantoinase; Bacillus dihydroorotase homology
C:Keywords: hydrolase

Query Match 70.2%; Score 40; DB 2; Length 489;
Best Local Similarity 63.6%; Pred. No. 9.2;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSVDSYVFKPF 11
Db 412 SSIDYVFKPF 422

RESULT 13
T13719
calo protein - fruit fly (Drosophila melanogaster) (fragment)
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: T13719
R:Xu, X.Z.; Wes, P.D.; Chen, H.; Li, H.S.; Yu, M.; Morgan, S.; Liu, Y.; Montell, C.
J. Biol. Chem. 273, 31297, 1998
A:Title: Retinal targets for calmodulin include proteins implicated in synaptic trans
A:Reference number: Z17709; MUID:99030403
A:Accession: T13719
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

A:Residues: 1-4116 <XUX>
A:Cross-references: EMBL:Y17920; NID:g3893108; PIDN:CAA76940.1; PID:g3893109
C:Genetics:
A:Gene: calo
A:Cross-references: Flybase:Fbgn0011230

Query Match 68.4%; Score 39; DB 2; Length 4116;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 DYSVFKPF 11
11:111111
DB 4013 DYTIVKPF 4020

RESULT 14
H82183
conserved hypothetical protein VC1567 [imported] - Vibrio cholerae (strain N16961 serog
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: H82183
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; M0ID:20406833
A:Accession: H82183
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-425 <HEI>
A:Cross-references: GB:AE004234; GB:AE003852; NID:g9656070; PIDN:AAF94721.1; GSPDB:GN001
C:Genetics:
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
A:Gene: VC1567
A:Map position: 1

Query Match 66.7%; Score 38; DB 2; Length 425;
Best Local Similarity 63.6%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SVDYSVFKPF 11
1:111111
DB 209 SPIDSGFTPF 219

RESULT 15
T37549
hypothetical protein SPAC1167.06c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37549
R:McLean, J.; Harris, D.; Bartell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL data library, September 1997
A:Reference number: Z21722
A:Accession: T37549
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-430 <MC1>
A:Cross-references: EMBL:Z99161; PIDN:CAB16211.1; GSPDB:GN00066; SPDB:SPAC1167.06c
C:Genetics:
A:Experimental source: strain 972h-; cosmid c1167
A:Gene: SPDB:SPAC1167.06c
A:Map position: 1

Query Match 66.7%; Score 38; DB 2; Length 430;
Best Local Similarity 70.0%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 SVDYSVFKPF 11

DB 21 SVDYSQFKAF 30
1:1111111

Search completed: June 13, 2002, 08:55:51
Job time: 371 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:53:26 ; Search time 158.68 Seconds
(without alignments)
4.900 Million cell updates/sec

Title: US-09-710-339-2_COPY_161_167

Sequence: 1 SSODYFH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	478	15	AA46065
2	40	100.0	478	15	AA46065
3	40	100.0	478	16	AA72450
4	40	100.0	478	16	AA72450
5	40	100.0	478	16	AA72450
6	40	100.0	478	16	AA72450
7	40	100.0	478	16	AA72450
8	40	100.0	478	16	AA72450
9	40	100.0	478	16	AA72450
10	40	100.0	478	16	AA72450
11	40	100.0	478	16	AA72450

12	33	82.5	124	22	AA05772	Human polypeptide
13	32	80.0	501	22	ABG15126	Novel human diapo
14	32	80.0	529	22	ABB70891	Drosophila melanog
15	32	80.0	583	18	AA26654	Human PANC1A assoc
16	32	80.0	594	21	AA85501	Human prostate can
17	32	80.0	700	20	AA736573	B. bassiana POPS r
18	32	80.0	715	20	AA79737	Human protein SEQ
19	32	80.0	1118	19	AAW82395	Human ubp protein
20	32	80.0	1128	22	ABG11843	Novel human diapo
21	32	80.0	1130	22	ABG2207	Human human diapo
22	31	77.5	69	20	AA67870	Human secreted pro
23	31	77.5	89	22	AA06897	Human polypeptide
24	31	77.5	104	21	AA90470	Yeast Sm11 (suppre
25	31	77.5	119	22	AA931764	Novel human secret
26	31	77.5	123	22	AA008298	Human polypeptide
27	31	77.5	191	22	AAW80917	Human haematologic
28	31	77.5	191	22	AAW81318	Human haematologic
29	31	77.5	191	22	AAW81390	Human haematologic
30	31	77.5	191	22	AAW81634	Human haematologic
31	31	77.5	191	22	AAW81672	Human haematologic
32	31	77.5	208	22	ABG59074	Drosophila melanog
33	31	77.5	275	22	AAU08723	Human clone h01143
34	31	77.5	275	22	AAW73980	Human starazin-11
35	31	77.5	343	21	AAW43328	Human cancer assoc
36	31	77.5	374	21	AAW43328	Human ORFX ORF2092
37	31	77.5	461	20	AAW35506	Chlamydia pneumoni
38	31	77.5	474	21	AAW56509	Human Jurkat cell
39	31	77.5	500	22	AAE10164	RNA helicase conse
40	31	77.5	614	22	ABB44571	Mouse wound healin
41	31	77.5	614	22	ABB44572	Human wound healin
42	31	77.5	615	22	ABB44573	Mouse wound healin
43	31	77.5	2427	21	AAW51843	Marine embryo macr
44	30	75.0	25	17	AAW01178	Serine protease Pf
45	30	75.0	25	22	AAW50574	Flea serine protea

ALIGNMENTS

RESULT 1	
AA46065	AA46065 standard; protein; 478 AA.
AC	AA46065;
DT	18-JUL-1994 (first entry)
XX	
DE	Mutant alpha-amylase.
KW	Methionine substitution; stability; activity; detergent;
KW	dishwashing agents; liquidation agents.
XX	
OS	Aspergillus oryzae.
XX	
PN	W09402597-A.
PD	03-FEB-1994.
XX	
PF	06-JUL-1993; 93WO-DK00230.
XX	
PR	23-JUL-1992; 92DK-0000946.
PR	16-DEC-1992; 92DK-0001503.
PR	15-MAR-1993; 93DK-0000292.
XX	
PA	(NOVO) NOVO-NORDISK AS.
XX	
PI	Bisgard-frantzen H, Svendsen A;
XX	
DR	WPI: 1994-048855/06.
PT	Mutant alpha-amylase from Bacillus species comprising a
PT	methionine substitution - with improved stability and activity at
PT	low pH, for use in detergents, dishwashing agents and

PF Liquefaction agents
XX
PS Claim 1; Page 7; 20pp; English.
XX
CC The sequence of that of the *Aspergillus oryzae* alpha amylase, sold
CC commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can
CC be mutated by substitution of one or more of its methionine residues
CC for any amino acid other than cysteine. The mutant alpha-amylase
CC exhibits a better activity level and better stability in the
CC presence of oxidizing agents than previous mutant alpha amylases,
CC and improved thermostability at moderately low pH. The enzyme can
CC be used as an additive for detergents, dishwashing agents and
CC liquifaction agents.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 40; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 141 ssqdyfh 147

RESULT 2
AAR79025 100.0%; Score 40; DB 15; Length 478;
ID AAR79025 standard; protein: 478 AA.
XX
AC AAR79025;
XX
DT 22-MAR-1996 (first entry)
XX
DE Mature taka-amylase A.
XX
KM Wild type; neopullulanase; B. stearotherophilus; mutant;
KM food industry; modification; hydrophobicity; replacement
KM insertion; deletion.
XX
OS *Aspergillus oryzae*.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 30..38
FT Disulfide-bond 150..164
FT Disulfide-bond 240..283
FT Disulfide-bond 439..474
XX
PN JP07177891-A.
XX
PD 18-JUL-1995.
XX
PF 31-OCT-1994; 94JP-0288658.
XX
PR 12-NOV-1993; 93JP-0306096.
XX
PA (EZAK) EZAKI GLICO CO.
PA (NIDE) NEC CORP.
XX
DR WPI; 1995-279919/37.
XX
PT Modifying a transferase by enhancing hydrophobicity of a selected
PT site - increases transfer activity, also new mutant
PT neo-pullulanase(s)
XX
PS Disclosure: Page 10-11; 18pp; English.
XX
CC This sequence represents the mature form of taka-amylase from A. *oryzae*.
CC This sequence was used in a method for the generation of mutant
CC pullulanases for use in the food industry (see also AAR79026-28). The
CC wild type pullulanase enzyme was modified by the method of the
CC invention for enhancing the hydrophobicity of a selected site of the
CC pullulanase. The method comprises replacement of a group in the

CC selected site with a hydrophobic group, replacement of an amino acid
CC with a hydrophobic amino acid, and/or insertion or deletion of a
CC hydrophobic amino acid from the selected site. The method was used
CC to produce neopullulanases Y37F, S422V and M375L.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 40; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 141 ssqdyfh 147

RESULT 3
AAR72450 100.0%; Score 40; DB 16; Length 478;
ID AAR72450 standard; protein: 478 AA.
XX
AC AAR72450;
XX
DT 01-DEC-1995 (first entry)
XX
DE *Aspergillus oryzae* alpha amylase (mature protein).
XX
KM Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
KM washing; *Bacillus licheniformis*; *Bacillus amyloliquefaciens*;
KM *Bacillus stearotherophilus*; dyeing; bleaching; scouring; textile;
KM thermostable.
XX
OS *Aspergillus oryzae*.
XX
PN W09510603-A.
XX
PD 20-APR-1995.
XX
PF 05-OCT-1994; 94WO-DK00370.
XX
PR 08-OCT-1993; 93DK-0001133.
PR 02-FEB-1994; 94DK-0000140.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Svendsen A, Theilersen M;
PI Van der zee P, Bisgaardfrantzen H, Borchert T;
XX
DR WPI; 1995-161790/21.
XX
PT New *Bacillus* derived alpha-amylase variants - having amino acid
PT modifications to improve washing and/or dishwashing performance
XX
PS Disclosure: Page 75-76; 105pp; English.
XX
CC Variant alpha amylase enzymes which have improved washing and/or
CC as detergent additives. The enzymes have one or more amino acid
CC residues added, deleted or substituted. The variants can also be
CC used for textile desizing prior to scouring, bleaching and dyeing.
CC The variants have improved thermostability, acid/alkaline stability;
CC low temperature optimum; pH optimum; higher hydrolysis velocity and
CC improved tolerance to other composition constituents, e.g. oxidation
CC agents.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 40; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSQDYFH 7
DB 141 ssqdyfh 147

FT			residue of a parent alpha-amylase (used as a template for a variant) corresponding to this fragment is deleted or replaced with a fragment corresponding to 102-199 of AAM14498; claim 42"
FT	Misc-difference	165..177	
FT	/label=	loop 3 modification region	
FT	/note=	"at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to 195-202 of AAM14499 is deleted or replaced with a fragment corresponding to this fragment; claim 21"	
FT	Misc-difference	166..173	
FT	/label=	loop 3 modification region	
FT	/note=	"preferred region where at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to 196-198 of AAM14499 is deleted or replaced with a fragment corresponding to this fragment; claim 23"	
FT	Misc-difference	291..313	
FT	/label=	loop 8 modification region	
FT	/note=	"at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to 322-346 of AAM14498 is deleted or replaced with a fragment corresponding to this fragment; claim 36"	
FT	Misc-difference	297..313	
FT	/label=	loop 8 modification region	
FT	/note=	"preferred region where at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to 323-345 of AAM14498 is deleted or replaced with a fragment corresponding to this fragment; claim 38"	
PN			
XX	WO9623874-A1.		
PD	08-AUG-1996.		
XX			
PR	05-FEB-1996;	96WO-DK00057.	
PR	10-NOV-1995;	95DK-0001256.	
PR	03-FEB-1995;	95DK-0000128.	
PR	23-OCT-1995;	95DK-0001192.	
PA	(NOVO) NOVO-NORDISK AS.		
XX			
PI	Bisgard-frantzen H, Borchert TV, Svendsen A;		
DR	WPI; 1996-371424/37.		
XX			
PT	Alpha-amylase variants and methods of production - have altered properties such as calcium dependency, substrate binding and stability		
PT			
PS	Disclosure; Page 87-88; 171pp: English.		
XX			
CC	The present sequence is the mature Aspergillus oryzae alpha-amylase (A.).		
CC	Variants of parent termamyl- and fungamyl-like alpha-amylases (and methods of constructing them) are claimed. Examples of variants are featured above. The variants have altered properties such as calcium dependency, substrate binding and stability. Also one or more proline or cysteine residues in the variant is modified or replaced with a non-proline or non-cysteine residue such as alanine. The variants can be used for (dis)washing, as detergent additives or for fabric desizing on starch liquefaction. They can also be used for the production of sweeteners and ethanol from starch. See also AAM14498-99.		
XX	Sequence	478 AA;	

	Query Match	Similarity	100.0%;	Score	40;	DB	17;	Length	478;
	Best Local	Similarity	100.0%;	Pred.	No.	6.7;			
Matches	7;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
OY	1 SSQDYFH 7								
Db	141 ssqdyfh 147								
	RESULT	6							
ID	AAB84206	standard; Protein:	498 AA.						
XX	AAB84206								
AC	AAB84206;								
XX	06-AUG-2001	(first entry)							
DY									
XX	Amino acid sequence of a fungamyl-like alpha-amylase.								
DE									
XX	Fungamyl-like alpha-amylase; glucoamylase; dextrinisation; maltose;								
KW	alcohol; starch; dough improver; brewing; starch liquification.								
XX									
OS	Aspergillus oryzae.								
PN	WO200134784-A1.								
PD	17-MAY-2001.								
XX									
PF	10-NOV-2000; 2000WO-DK00626.								
PR	10-NOV-1999; 99DK-0001617.								
XX									
PA	(NOVO) NOVOZYMES AS.								
XX									
PI	Bisgard-Frantzen H, Svendsen A, Pedersen S;								
DR	WPI; 2001-367478/38.								
N-P	PSDB: AAF90208.								
PT	New variant of Fungamyl-like alpha-amylase, useful for production of								
PT	maltose syrups. Includes mutations that improve stability against heat								
PT	and acidic pH -								
XX									
PS	Claim 1; Page 42-45; 49pp; English.								
XX									
CC	The present sequence represents a fungamyl-like alpha-amylase. The								
CC	specification describes variants of this fungamyl-like alpha-amylase,								
CC	which have an alteration in one the amino acid regions 98-110, 150-160,								
CC	161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion								
CC	or substitution of an amino acid or an insertion of an amino acid								
CC	downstream of a particular position. The variants retain alpha-amylase								
CC	activity, and have better heat stability and/or stability at acidic pH,								
CC	relative to wild-type enzyme. The variants can therefore be used at								
CC	higher temperatures (more efficient conversion or faster reaction), and								
CC	have reduced need for cooling and reduced risk of contamination). The								
CC	variants may also be used in conjunction with other enzymes,								
CC	particularly glucamylase during dextrinisation. The variants are								
CC	used to produce syrups, particularly of high maltose content, or alcohol,								
CC	from starch, as dough improver for baked goods; in brewing, to increase								
CC	fermentability of the wort; and for liquefaction of starch.								
XX									
SQ	Sequence	498 AA;							
	Query Match	100.0%;	Score	40;	DB	22;	Length	498;	
	Best Local	Similarity	100.0%;	Pred.	No.	6.9;			
Matches	7;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
OY	1 SSQDYFH 7								
Db	161 ssqdyfh 167								


```

RESULT 7
AA088212
ID AAR88212 standard; Protein; 493 AA.
XX
AC AAR88212:
XX
DT 03-APR-1996 (first entry)
XX
DE Alpha-amylase.
XX
KW Alpha-amylase; thermostable enzyme; baking; Aspergillus niger.
XX
OS Thermomyces lanuginosus CBS 224.63.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT /label= Sig_peptide
XX
PN WO9601323-A1.
XX
PD 18-JAN-1996.
XX
PF 03-JUL-1995; 95WO-EP02607.
XX
PR 04-JUL-1994; 94GB-0013419.
XX
PA (DANT-) DANISCO AS.
XX
PI Michelsen B, Rasmussen P;
XX
DR WPI: 1996-087673/09.
DR N-PSDB; AAT10562.
XX
PT Thermophilic alpha-amylase with activity range of 60-80 degrees C -
PT derived from Thermomyces lanuginosus, useful in the prepn. of
PT foodstuffs and bakery prods. esp. bread
XX
PS Claim 3; Page 36-38; 94pp; English.
XX
CC A thermostable alpha-amylase (AAR88212) of Thermomyces lanuginosus
CC CBS 224.63 is expressed in Aspergillus niger hosts using a gene
CC (AAT10562) isolated from a T. lanuginosus gene library. The
CC recombinant enzyme (54-60 kDa) shows optimal activity at
CC 60-70 deg and pH 5.8-6, has a pi of 3.7 and is active at 60-80
CC deg.
XX
SQ Sequence 493 AA;

```

```

Query Match 92.5%; Score 37; DB 17; Length 493;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSODYFH 7
DB 161 nsqdyfh 167

```

```

RESULT 8
AA008447
ID AAO08447 standard; Protein; 52 AA.
XX
AC AAO08447;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 22339.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.

```

```

XX Homo sapiens.
OS
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Dymnac RT;
XX
DR WPI: 2001-514838/56.
DR N-PSDB; AAI88378.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 20; SEQ ID NO 22339; 1399pp + Sequence listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 52 AA;

```

```

Query Match 85.0%; Score 34; DB 22; Length 52;
Best Local Similarity 85.7%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SSODYFH 7
DB 2 ssqdyhm 8

```

```

RESULT 9
AAG36453
ID AAG36453 standard; Protein; 1222 AA.
XX
AC AAG36453;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 44675.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PP 25-FEB-2000; 2000EP-0301439.

```

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137722.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.0%; Score 34; DB 21; Length 1222;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SSQDYFH 7
111111
Db 1130 ssadyfh 1136

RESULT 10

AAG36452
ID AAG36452 standard; Protein; 1257 AA.

AC AAG36452;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 44674.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PM EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0135621.
PR 27-MAY-1999; 99US-0135629.
PR 28-MAY-1999; 99US-0135782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144634.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144815.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-015138.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157533.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.0%; Score 34; DB 21; Length 1557;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSQDYPH 7
111111
Db 1165 ssadyth 1171

RESULT 11
AAG36451
ID AAG36451 standard; Protein; 1275 AA.
XX
AC AAG36451;
XX

DT 18-OCT-2000 (first entry)
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 44673.
XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.
XX

PN EP1033405-A2.
XX

PD 06-SEP-2000.
XX

PE 25-FEB-2000; 2000EP-0301439.
XX

XX 25-FEB-1999; 99US-0121825.
XX

PR 05-MAR-1999; 99US-0123180.
PR

PR 09-MAR-1999; 99US-0123548.
PR

PR 23-MAR-1999; 99US-0125788.
PR

PR 25-MAR-1999; 99US-0126264.
PR

PR 29-MAR-1999; 99US-0126785.
PR

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.0%; Score 34; DB 21; Length 1275;
Best Local Similarity 85.7%; Pred. No. 2; De+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SDDYFH 7
|||:|
Db 1183 ssadyfh 1189

RESULT 12
AA005772
ID AA005772 standard; Protein: 124 AA.

AC AA005772;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 19664.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KM vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
KM tissue growth factor; immunomodulatory; cancer; leukemia;
KM nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.

XX N-PSDB; AA185703.

PT Isolated nucleic acids and polypeptides, useful for preventing

PT diagnosing and treating e.g. leukemia, inflammation and immune

XX disorders -

CC Claim 20; SEQ ID NO 19664; 1399pp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AA179941-AA193841) and

CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, hematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 124 AA:

Query Match 82.5%; Score 33; DB 22; Length 124;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 SDDYFH 7
|||:|
Db 101 sqdyfh 106

RESULT 13
ABG15126
ID ABG15126 standard; Protein: 501 AA.

AC ABG15126;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #15117.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KM food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

XX N-PSDB; AA579313.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

CC Claim 20; SEQ ID NO 45485; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 CC Sequence 501 AA:

Query Match 80.0%; Score 32; DB 22; Length 501;
 Best Local Similarity 71.4%; Pred. No. 2.6e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 111111
 Db 410 smdyfh 416

RESULT 14
 ID ABB70891
 XX ABB70891 standard; Protein; 529 AA.

AC ABB70891;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 39465.

XX Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW.

DR WPI; 2001-656860/75.

XX N-PSDB; ABL14994.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Disclosure: SEQ ID NO 39465; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABU01840-ABU16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 529 AA:

Query Match 80.0%; Score 32; DB 22; Length 529;
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 111111
 Db 389 ssqdyfh 395

RESULT 15
 ID AAW26654
 XX AAW26654 standard; Protein; 583 AA.

AC AAW26654;

DT 16-FEB-1998 (first entry)

XX Human PANC1A associated with pancreatic cancer.

XX PANC1A; pancreas cancer; diagnosis; therapy; human.

XX Homo sapiens.

PN WO9724437-A1.

PD 10-JUL-1997.

PF 18-DEC-1996; 96WO-US19981.

XX 15-MAR-1996; 96US-0616392.

PR 29-DEC-1995; 95US-0581240.

XX (INCY-) INCYTE PHARM INC.

XX Akerblom IE, Coleman R, Gould RD, Seilhamer JJ;

DR WPI; 1997-363673/33.

XX N-PSDB; AAT91010.

XX DNA encoding PANC1A and PANC1B associated with pancreatic cancer -
 PT useful for diagnosis of pancreatic cancer and identifying risk of
 PT cancer

XX Disclosure: Page 46-47; 65pp; English.

XX This protein comprises human PANC1A, which is associated with
 CC pancreatic cancer and possibly other cancers. Its sequence was
 CC deduced from a full-length PANC1A cDNA clone (see AAT91010) derived
 CC from human teratocarcinoma hMT-2 cells. Specific fragments of
 CC PANC1A (see AAT91008) and PANC1B (see AAT91009) cDNA are claimed.
 CC Also claimed are: antisense molecules; expression vectors;
 CC transformed host cells; and purified PANC1A and PANC1B
 CC polypeptides. Diagnostic testing for the presence of PANC1A or
 CC PANC1B in a sample provides a means for positively correlating
 CC abnormal levels of PANC1A or PANC1B with pancreatic cancer
 CC (claimed). The host cells can be cultured for the recombinant
 CC production of PANC1A or PANC1B proteins (claimed). PANC1A and
 CC PANC1B polypeptides can be used to screen for antagonists and
 CC inhibitors, and to raise diagnostic antibodies.

XX Sequence 583 AA:

Query Match 80.0%; Score 32; DB 18; Length 583;
 Best Local Similarity 71.4%; Pred. No. 3.1e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 111111
 Db 433 stdyfh 439

Fri Jun 14 10:36:43 2002

us-09-710-339-2_copy_161_167.rag

Page 12

Search completed: June 13, 2002, 08:53:28
Job time: 278 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:55:51 ; Search time 73.31 Seconds
(without alignments)
9.175 Million cell updates/sec

Title: us-09-710-339-2_COPY_161_167

Perfect score: 40

Sequence: 1 SSODYFH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length DB	ID	Description
1	40	100.0	478	2 JK0201	alpha-amylase (EC
2	40	100.0	498	2 A48305	alpha-amylase (EC
3	40	100.0	499	1 ALAS1	alpha-amylase (EC
4	40	100.0	499	1 ALAS3	alpha-amylase (EC
5	40	100.0	499	2 B48305	alpha-amylase (EC
6	40	100.0	499	2 JS0653	alpha-amylase (EC
7	40	100.0	499	2 JT0466	alpha-amylase (EC
8	40	100.0	499	2 JN0588	alpha-amylase (EC
9	40	100.0	624	1 JC4510	alpha-amylase (EC
10	34	85.0	130	2 H71415	hypothetical prote
11	33	82.5	420	2 T42616	envelope protein -
12	33	82.5	601	2 T18767	hypothetical prote
13	33	82.5	1131	2 T41943	major DNA binding
14	33	82.5	4351	2 T00252	MEGF1 protein - ra
15	32	80.0	144	2 JC7121	hypothetical gland h
16	32	80.0	180	2 B95039	hypothetical prote
17	32	80.0	289	2 A43552	homeotic protein H
18	32	80.0	317	2 E84088	restiction endonu
19	32	80.0	351	2 E86187	YJP8H12.10 (import
20	32	80.0	411	2 S61245	hypothetical prote
21	32	80.0	419	2 T15199	conserved hypochet
22	32	80.0	482	2 A90248	probable membrane
23	32	80.0	685	2 S67146	thra bifunctional
24	32	80.0	820	2 D56CK	aspartokinase I-ho
25	32	80.0	820	2 B90629	aspartokinase I-ho
26	32	80.0	820	2 B85480	aspartokinase I-ho
27	32	80.0	820	2 AC0502	aspartokinase I-ho
28	31	77.5	67	2 T17658	hypothetical prote
29	31	77.5	104	2 S49803	hypothetical prote

30	31	77.5	214	2 B75513	pyridoxamine 5-pho
31	31	77.5	261	2 AF1307	hypothetical prote
32	31	77.5	261	2 AF1679	hypothetical prote
33	31	77.5	386	2 S72435	RNA-binding protei
34	31	77.5	461	2 F86597	UDP-glucose pyroph
35	31	77.5	461	2 H72024	UDP-N-acetylglucos
36	31	77.5	597	2 T35746	hypothetical prote
37	31	77.5	614	1 I48385	hypothetical prote
38	31	77.5	614	2 JC1087	RNA helicase TN22
39	31	77.5	668	2 T10575	RNA helicase, ATP-
40	31	77.5	825	2 S55060	hypothetical prote
41	31	77.5	893	2 T03864	fertilin alpha-II
42	31	77.5	905	2 S55059	hypothetical prote
43	30	75.0	182	2 AE2146	fertilin alpha-I
44	30	75.0	194	2 T22209	hypothetical prote
45	30	75.0	229	2 A89473	hypothetical prote

ALIGNMENTS

RESULT 1

JK0201 alpha-amylase (EC 3.2.1.1) - Aspergillus oryzae

N:Alternate names: glycogenase; Taka-amylase A

C:Species: Aspergillus oryzae

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998

C:Accession: JK0201

R: Toda, H.; Kondo, K.; Narita, K.

Proc. Jpn. Acad. 58B, 208-212, 1982

A:Title: The complete amino acid sequence of Taka-amylase A.

A:Reference number: JK0201

A:Accession: JK0201

A:Molecule type: protein

A:Residues: 1-478 <TOD>

C:Comment: One atom of calcium per molecule is essential for the activity.

C:Comment: This enzyme is a glycoprotein.

C:Comment: See also PIR:JT0466 and PIR:JS0240.

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Species: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:173-300/Domain: alpha-amylase core homology <AMY>

F:197/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 40; DB 2; Length 478;

Best Local Similarity 100.0%; Pred. No. 1.3;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 141 SSODYFH 147

RESULT 2

A48305 alpha-amylase (EC 3.2.1.1) A precursor - Aspergillus awamori

C:Species: Aspergillus awamori

C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997

C:Accession: A48305

R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J

Curr. Genet. 17, 203-212, 1990

A:Title: Cloning, characterization, and expression of two alpha-amylase genes from As

A:Reference number: A48305; MUID:90254827

A:Accession: A48305

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-498 <KOR>

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycoen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 40; DB 2; Length 496;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSQDYFH 7
 Db 162 SSQDYFH 168

RESULT 3

ALAS1

alpha-amylase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae
 N:Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae
 C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 15-Sep-2000
 C:Accession: S04548; A33214; J50240; A91930; A93767; A10627
 R:Mirsel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989
 A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon
 A:Reference number: S04548; MUID:89237897
 A:Accession: S04548
 A:Molecule type: DNA
 A:Residues: 1-499 <WIR>
 A:Cross-references: EMBL:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921
 A:Genetics: AMY1
 A:Accession: A33214
 A:Molecule type: mRNA
 A:Residues: 1-499 <WIR>
 A:Cross-references: GB:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921
 R:Genes, M.J.; Dove, M.J.; Seligy, V.L.
 Gene 79, 107-117, 1989
 A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
 A:Reference number: J50240; MUID:89378767
 A:Accession: J50240
 A:Molecule type: DNA
 A:Residues: 1-499 <GEN>
 A:Genetics: AMY2
 A>Note: the authors refer to this as isozyme II
 R:Riemura, S.; Ikenaka, T.
 J. Biochem. 74, 1-10, 1973
 A:Reference number: A91930; MUID:74001521
 A:Accession: A91930
 A:Molecule type: protein
 A:Residues: 206-225 <ISE>
 R:Narita, K.
 Proc. Jpn. Acad. 51, 285-290, 1975
 A:Reference number: A93767
 A:Accession: A93767
 A:Molecule type: protein
 A:Residues: 434-443, 446-447, 'Q', 449-458, 'GTY', 459-464, 467-468, 'B', 470, 'B', 472-499 <NAR>
 R:Matsura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
 J. Biochem. 95, 697-702, 1984
 A:Title: Structure and possible catalytic residues of Taka-amylase A.
 A:Reference number: A37454; MUID:84212370
 A:Accession: A37454
 A:Contents: annotation: X-ray crystallography, 3.0 angstroms
 R:Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkenburg, J.P.; Wilkinson, A.
 submitted to the Brookhaven Protein Data Bank, August 1992
 A:Reference number: A51548; PDB:6F7AA
 A:Contents: annotation: X-ray crystallography, 2.1 angstroms, residues 22-497
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics: <AMY1>
 A:Gene: amy1
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Genetics: <AMY2>
 A:Gene: amy2
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycoen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; po
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase 1 #status experimental <NAT>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 40; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSQDYFH 7
 Db 162 SSQDYFH 168

RESULT 4

ALAS3

alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae
 N:Alternate names: alpha-amylase isozyme I; glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae
 C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
 C:Accession: S04549; A33215; A44713
 R:Mirsel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989
 A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exo
 A:Reference number: S04549; MUID:89237897
 A:Accession: S04549
 A:Molecule type: DNA
 A:Residues: 1-499 <WIR>
 A:Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 A:Accession: A33215
 A:Molecule type: translation not shown
 A:Residues: 1-499 <WIR>
 A:Cross-references: GB:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 R:Genes, M.J.; Dove, M.J.; Seligy, V.L.
 Gene 79, 107-117, 1989
 A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each contain
 A:Reference number: J50240; MUID:89378767
 A:Accession: A44713
 A:Molecule type: DNA
 A:Residues: 1-499 <GEN>
 A:Note: the authors refer to this as isozyme I
 R:Matsura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
 J. Biochem. 95, 697-702, 1984
 A:Title: Structure and possible catalytic residues of Taka-amylase A.
 A:Reference number: A37454; MUID:84212370
 A:Accession: A37454
 A:Contents: annotation: X-ray crystallography, 3.0 angstroms
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics: <AMY1>
 A:Gene: amy1
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycoen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase 3 #status experimental <NAT>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 40; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
|||||
Db 162 SSODYFH 168

RESULT 5
B48305
alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori

C:Species: Aspergillus awamori
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: B48305
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus
A:Reference number: A48305; MUID:90254827
A:Accession: B48305
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499 <KOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
C:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>
F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 40; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
|||||
Db 162 SSODYFH 168

RESULT 6

JS0663
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
C:Species: Aspergillus sp.
C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
C:Accession: JS0663
R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
Biosci. Biotechnol. Biochem. 56, 174-179, 1992
A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shiroyamii and its expression
A:Reference number: JS0663; MUID:92323146
A:Accession: JS0663
A:Molecule type: mRNA
A:Residues: 1-499 <SH1>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-31/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <AMY>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 40; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
|||||
Db 162 SSODYFH 168

RESULT 7
JT0466
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae

N:Alternate names: glycosenase; Taka-amylase A
C:Species: Aspergillus oryzae
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: JT0466
R:Tada, S.; Jimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
Agric. Biol. Chem. 53, 593-599, 1989
A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Aspergillus
A:Reference number: JT0466
A:Accession: JT0466
A:Molecule type: DNA
A:Residues: 1-499 <TAD>
C:Comment: See also PIR:JK0201 and PIR:JS0240.
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics:
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <AMY>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 40; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
|||||
Db 162 SSODYFH 168

RESULT 8

JN0588
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: Taka-amylase A
C:Species: Aspergillus oryzae
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
C:Accession: JN0588
R:Tsuikagoshi, N.; Furukawa, M.; Nagaba, H.; Kirita, N.; Teubol, A.; Ueda, S.
Gene 84, 319-327, 1989
A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for
A:Reference number: JN0588; MUID:90128276
A:Accession: JN0588
A:Molecule type: mRNA
A:Residues: 1-499 <TSU>
C:Comment: The alpha amylases are encoded by multigene family.
C:Genetics:
A:Gene: Taa-G1
A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <AMY>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 40; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7

Db 162 SSODYFH 168
|||||||

RESULT 9

alpha-amylase (EC 3.2.1.1) precursor - yeast (Lipomyces kononenkoae)
N:Alternate names: LKAI protein; raw starch-degrading amylase
N:Contains: alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41)
C:Species: Lipomyces kononenkoae
C:Date: 07-Feb-1996 #sequence_revision 23-Aug-1996 #text_change 18-Jun-1999
C:Accession: J04510; PC4116
R:Steyn, A.J.C.; Marmur, J.; Pretorius, I.S.
Gene 160, 65-71, 1995
A>Title: Cloning, sequence analysis and expression in yeasts of a cDNA containing a lip
A:Reference number: J04510; M01D:96105202
A:Accession: J04510
A:Molecule type: mRNA
A:Residues: 1-624 <STR>
A:Cross-references: GB:U30376; NID:q1173536; PIDN:AA049622.1; PID:q1173537
A:Experimental source: strain IG04052B
A:Accession: PC4116
A:Molecule type: protein
A:Residues: 29-44 <STR>
A:Experimental source: IG04052B
C:Genetics:
A:Gene: LKAI
C:Function:
A:Description: hydrolyzes raw starch, degrading both alpha-1,4 and alpha-1,6 linkages
A:Pathway: glycogen/starch degradation
C:Superfamily: Lipomyces alpha-amylase; alpha-amylase core homology; glucosylase starch
C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-624/Product: alpha-amylase #status predicted <MAT>
F:48-141/Domain: glucosylase starch-binding domain homology <SBD>
F:320-447/Domain: alpha-amylase core homology <AM>
F:177-185,287-311,387-430,587-622/Disulfide bonds: #status predicted
F:304,344/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:357,377,444/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 40; DB 1; Length 624;
Best Local Similarity 100.0%; Pred. No. 1.8; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSODYFH 7

|||||||
Db 288 SSODYFH 294

RESULT 10

hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C:Accession: H71415
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambolt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, S.; Ertl, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Pulgomech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palmer, K.; Benes, V.; Reichman, S.; Ans
C.; Chalvatzis, N.
A>Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis tha
A:Reference number: A71400; M01D:98121113
A:Accession: H71415
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-130 <BEV>
A:Cross-references: GB:297338; NID:q2244870; PID:q2244877
C:Genetics:
A:Map position: 4COP9-4G3845

Query Match 85.0%; Score 34; DB 2; Length 130;
Best Local Similarity 85.7%; Pred. No. 5.6; Mismatches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SSODYFH 7

|||||||
Db 38 SSADYFH 44

RESULT 11

envelope protein - equine herpesvirus 4 (strain NS80567)
C:Species: equine herpesvirus 4
A:Variety: strain NS80567
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T42616
R:Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.
J. Gen. Virol. 79, 1197-1203, 1998
A>Title: The DNA sequence of equine herpesvirus-4.
A:Reference number: T22173; M01D:98284497
A:Accession: T42616
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-420 <TEB>
A:Cross-references: EMBL:AF030027; NID:q2605950; PIDN:AA059593.1; PID:q2606021
A:Experimental source: strain NS80567
C:Genetics:
A:Note: 73
C:Superfamily: pseudorabies virus glycoprotein gp63

Query Match 82.5%; Score 33; DB 2; Length 420;
Best Local Similarity 83.3%; Pred. No. 34; Mismatches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SODYFH 7

|||||||
Db 114 NODYFH 119

RESULT 12

hypothetical protein BE10.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18787
R:McMurray, A.
Submitted to the EMBL Data Library, March 1997
A:Reference number: Z19022
A:Accession: T18787
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-601 <WIL>
A:Cross-references: EMBL:Z03372; PIDN:CAB07545.1; GSPDB:GN00021; CESP:BE10.3
A:Experimental source: clone BE10
C:Genetics:
A:Gene: CESP:BE10.3
A:Map position: 3
A:Introns: 21/3; 54/3; 117/2; 157/1; 212/2; 533/3

Query Match 82.5%; Score 33; DB 2; Length 601;
Best Local Similarity 71.4%; Pred. No. 50; Mismatches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSODYFH 7

|||||||
Db 453 NSODYFH 459

RESULT 13

T41943
 major DNA binding protein - human herpesvirus 7 (strain J1)
 C:Species: human herpesvirus 7
 C:Variate: strain J1
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T41943
 R:Nicholas, J.
 submitted to the EMBL Data Library, December 1995
 A:Description: Determination and analysis of the complete nucleotide sequence of human h
 A:Reference number: 222022
 A:Accession: T41943
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1131 <N1C>
 A:Cross-references: EMBL:U43400; PIDN:AAC54703.1
 A:Experimental source: strain J1
 C:Genetics:
 A:Note: U41
 C:Superfamily: herpesvirus DNA-binding protein

Query Match 82.5%; Score 33; DB 2; Length 1131;
 Best Local Similarity 71.4%; Pred. No. 1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 :|||
 Db 943 ASRDYFH 949

RESULT 14
 T00252
 MEGF1 protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
 C:Accession: T00252
 R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
 Genomics 51, 27-34, 1998
 A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
 A:Reference number: 214126; MUID:98360089
 A:Accession: T00252
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-4351 <NAK>
 A:Cross-references: EMBL:AB011527; NID:93449285; PIDN:BA032458.1; PID:93449286
 A:Experimental source: brain
 C:Genetics:
 A:Gene: MEGF1
 C:Superfamily: rat MEGF1 protein; cadherin repeat homology; EGF homology; laminin G rep
 F:3758-3949/Domain: laminin G repeat homology <LGR>
 F:3953-3985/Domain: EGF homology <EGF>
 F:3992-4023/Domain: EGF homology <EGF1>

Query Match 82.5%; Score 33; DB 2; Length 4351;
 Best Local Similarity 71.4%; Pred. No. 4.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 :|||
 Db 605 NEODYFH 611

RESULT 15
 JC7121
 androgenic gland hormone precursor - common pill bug
 C:Species: Armadillidium vulgare (common pill bug)
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
 C:Accession: JC7121
 R:Okuno, A.; Hasegawa, Y.; Ohira, T.; Katakura, Y.; Nagasawa, H.
 Biochem. Biophys. Res. Commun. 264, 419-423, 1999
 A:Title: Characterization and cDNA cloning of androgenic gland hormone of the terrestria
 A:Reference number: JC7121; MUID:20001935

A:Accession: JC7121
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-144 <OKO>
 A:Cross-references: DDBJ:AB029615; GB:AB029615; NID:96446571; PID:96446572
 A:Experimental source: androgenic gland

Query Match 80.0%; Score 32; DB 2; Length 144;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 QDYFH 7
 :|||
 Db 82 QDYFH 86

Search completed: June 13, 2002, 08:55:53
 Job time: 373 sec

Fri Jun 14 10:36:45 2002

us-09-710-339-2_copy_161_167.rpr

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:51 ; search time 38.1 Seconds
(without alignments)
7.114 Million cell updates/sec

Title: US-09-710-339-2_COPY_161_167

Perfect score: 40

Sequence: 1 SSQDYFH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	40	100.0	498 1	AMVA_ASPAW
2	40	100.0	499 1	AMVA_ASPOR
3	40	100.0	499 1	AMVA_ASPAM
4	40	100.0	499 1	AMVA_ASPSH
5	33	82.5	1131 1	DNBI_HSV7J
6	32	80.0	144 1	AGH_ARMVU
7	32	80.0	289 1	HXDB_MOUSE
8	32	80.0	411 1	VGLM_HSVBC
9	32	80.0	820 1	AKIH_ECOLI
10	32	80.0	1118 1	UBP8_HUMAN
11	31	77.5	104 1	YME8_YEAST
12	31	77.5	198 1	CCG7_HUMAN
13	31	77.5	208 1	LBW_DROME
14	31	77.5	386 1	TIAL_MOUSE
15	31	77.5	614 1	DDX5_HUMAN
16	31	77.5	614 1	DDX5_MOUSE
17	30	75.0	346 1	YDGT_MOUSE
18	30	75.0	432 1	YRKO_ECOLI
19	30	75.0	623 1	YRKO_BACSU
20	30	75.0	650 1	RSD1_YEAST
21	30	75.0	714 1	FRDA_HELPJ
22	30	75.0	714 1	FRDA_HELPY
23	30	75.0	937 1	CS32_ECOLI
24	30	75.0	1003 1	SYG_CHLTR
25	29	72.5	144 1	YITE_YEAST
26	29	72.5	217 1	THIE_HELPJ
27	29	72.5	219 1	THIE_HELPY
28	29	72.5	231 1	ALSE_ECOLI
29	29	72.5	242 1	MTGA_KLEPN
30	29	72.5	248 1	PCRB_METHH
31	29	72.5	256 1	ERGB_NEUCR
32	29	72.5	312 1	NRT1_CHICK
33	29	72.5	374 1	CKR6_HUMAN

34	29	72.5	445 1	YG2V_YEAST
35	29	72.5	675 1	PPK_HELPJ
36	29	72.5	675 1	PPK_HELPY
37	29	72.5	953 1	Y442_HUMAN
38	29	72.5	1101 1	DIAD_HUMAN
39	29	72.5	1564 1	PDRA_YEAST
40	29	72.5	1584 1	YJ9G_YEAST
41	28	70.0	106 1	TPIS_STPMU
42	28	70.0	183 1	AAOI_DICDI
43	28	70.0	199 1	TDX_TRYBR
44	28	70.0	200 1	ATKC_ANASL
45	28	70.0	217 1	HNT2_YEAST

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	498 AA.
AMVA_ASPAW				
ID	AMVA_ASPAW			
AC	002905;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Alpha-amylose A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase A).			
DE	glucanohydrolase A).			
GN	AMVA.			
OS	Aspergillus awamori.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
OX	NCBI_TaxID=105351;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=OVK143F;			
RX	MEDLINE=90234627; PubMed=2340591;			
RA	Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,			
RA	Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,			
RA	Beika R.M.;			
RT	*Cloning, characterization, and expression of two alpha-amylose genes from Aspergillus niger var. awamori.;			
RT	Curr. Genet. 17:203-212(1990).			
RL	-1- CARBOLYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides.			
CC	-1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: X52755; CAA36966.1; -.			
DR	HSSP: P10529; 77NA.			
DR	InterPro: IPR00461; Alpha_Amylase.			
DR	Pfam: PF00128; alpha-amylose; 1.			
KW	Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;			
KW	Glycoprotein; Signal; Multigene family.			
FT	SIGNAL	1	21	POTENTIAL.
FT	CHAIN	22	498	ALPHA-AMYLASE A.
FT	ACT_SITE	227	227	BY SIMILARITY.
FT	ACT_SITE	231	231	BY SIMILARITY.
FT	ACT_SITE	251	251	BY SIMILARITY.
FT	ACT_SITE	318	318	BY SIMILARITY.
FT	DISULFID	51	59	BY SIMILARITY.
FT	DISULFID	171	185	BY SIMILARITY.
FT	DISULFID	261	304	BY SIMILARITY.
FT	DISULFID	461	496	BY SIMILARITY.
FT	CARBOHYD	218	218	N-LINKED (GLCNAC...) (POTENTIAL).
SQ	SEQUENCE	498 AA:	34880 MW:	763851IBC01ABA01 CRC64;

Query Match 100.0%; Score 40; DB 1; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.72; Mismatches 7; Conservative 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Indels 0; Gaps 0;

Oy 1 SSODYEH 7
 Db 162 SSODYEH 168

RESULT 2
 AMYA_ASPOK STANDARD; PRT; 499 AA.
 AC P10529; P11763; Q00250;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
 DE (1,4-alpha-D-glucan glucanohydrolase).
 GN AMY1 AND AMY2 AND AMY3.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 63503;
 RA Mitsel S., Lachmund A., Wildhardt G., Rutkowski E.;
 RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
 RT intron-exon organization.";
 RL Mol. Microbiol. 3:3-14(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89378767; PubMed=2789162;
 RA Genes M.J., Dove M.J., Selligy V.L.;
 RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
 RT containing eight introns.";
 RL Gene 79:107-117(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
 RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
 RT of Aspergillus oryzae.";
 RL Agric. Biol. Chem. 53:593-599(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90128276; PubMed=2612911;
 RA Tsunagoshi N., Furukawa M., Nagaba H., Kiritu N., Tsutol A.,
 RA Uekura S.;
 RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
 RT evidence for multiple related genes.";
 RL Gene 84:319-327(1989).
 RN [5]
 RP SEQUENCE OF 22-499.
 RA Toda H., Kondo K., Narita K.;
 RT "The complete amino acid sequence of Taka-amylase A";
 RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
 RN [6]
 RP SEQUENCE OF 206-225.
 RX MEDLINE=74001521; PubMed=4733850;
 RA Isemura S., Ikenaka T.;
 RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
 RT A with trypsin and chymotrypsin.";
 RL J. Biochem. 74:1-10(1973).
 RN [7]
 RP SEQUENCE OF 433-499.
 RA Narita K.;
 RT Proc. Jpn. Acad., B, Phys. Biol. Sci. 51:285-290(1975).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=80227691; PubMed=6156152;

RA Matsura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
 RA Toda H., Narita K., Kakudo M.;
 RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
 RT 3-A resolution.";
 RL J. Biochem. 87:1555-1558(1980).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=84212370; PubMed=6609921;
 RA Matsura Y., Kusunoki M., Harada W., Kakudo M.;
 RT "Structure and possible catalytic residues of Taka-amylase A.";
 RL J. Biochem. 95:697-702(1984).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
 RX MEDLINE=97428212; PubMed=9283074;
 RA Brzozowski A.M., Davies G.J.;
 RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
 RT inhibitor acarbose at 2.0-A resolution.";
 RL Biochemistry 36:10837-10845(1997).
 CC -I- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -I- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -I- MISCELLANEOUS: THE SEQUENCE OF AMY1 AND AMY2 IS SHOWN.
 CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X12725; CAA31218.1; -;
 DR EMBL: X12726; CAA31219.1; -;
 DR EMBL: X12727; CAA31220.1; -;
 DR EMBL: D00434; BAA00336.1; -;
 DR EMBL: M33218; AAA32708.1; -;
 DR PIR: S04548; ALAS1.
 DR PIR: S04549; ALAS3.
 DR PIR: A10627; A10627.
 DR PIR: JK0201; JK0201.
 DR PIR: JS0240; JS0240.
 DR PIR: J10466; J10466.
 DR PDB: 2TAA; 15-OCT-89.
 DR PDB: 6TAA; 31-OCT-93.
 DR PDB: 7TAA; 25-NOV-98.
 DR GLYCOSULEDB: P10529; -;
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase-1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 FT
 FT VARIANT 56 56 N-LINKED (GLCNAC. .).
 FT VARIANT 172 172 Q -> R (IN AMY3).
 FT
 FT CONFLICT 93 94 P -> L (IN AMY3).
 FT CONFLICT 106 106 TT -> DC (IN REF. 5).
 FT CONFLICT 184 184 Q -> T (IN REF. 5).
 FT CONFLICT 184 184 D -> Y (IN REF. 3).
 FT CONFLICT 195 195 P -> L (IN REF. 3).
 FT CONFLICT 255 255 G -> V (IN REF. 3).
 FT CONFLICT 291 291 D -> H (IN REF. 4).
 FT CONFLICT 345 345 I -> L (IN REF. 5).


```

FT CONFLICT 370 370 L -> A (IN REF. 4).
FT CONFLICT 406 409 WP1Y -> PYI (IN REF. 5).
FT CONFLICT 448 448 G -> S (IN REF. 5).
FT CONFLICT 497 497 S -> SD (IN REF. 5 AND 7).
FT HELIX 24 27
FT TURN 28 29
FT STRAND 32 35
FT HELIX 37 40
FT STRAND 42 42
FT TURN 43 44
FT HELIX 53 55
FT STRAND 61 61
FT HELIX 63 67
FT TURN 68 69
FT HELIX 70 74
FT TURN 75 77
FT STRAND 80 83
FT STRAND 87 89
FT STRAND 94 94
FT TURN 95 96
FT STRAND 97 97
FT TURN 100 101
FT STRAND 105 111
FT TURN 113 115
FT TURN 118 119
FT HELIX 120 130
FT TURN 131 132
FT STRAND 134 139
FT STRAND 143 143
FT STRAND 146 146
FT HELIX 150 152
FT TURN 155 156
FT STRAND 159 159
FT HELIX 164 166
FT STRAND 167 167
FT STRAND 172 172
FT HELIX 179 184
FT STRAND 186 188
FT STRAND 193 194
FT STRAND 196 197
FT TURN 199 200
FT HELIX 202 219
FT TURN 220 220
FT STRAND 223 226
FT TURN 227 228
FT HELIX 229 231
FT TURN 234 235
FT HELIX 236 244
FT STRAND 247 250
FT HELIX 257 265
FT TURN 266 266
FT STRAND 269 271
FT HELIX 273 283
FT TURN 286 287
FT HELIX 290 303
FT HELIX 307 309
FT STRAND 311 312
FT TURN 317 318
FT HELIX 322 325
FT HELIX 329 341
FT STRAND 345 349
FT TURN 350 351
FT HELIX 352 354
FT TURN 355 355
FT TURN 361 364
FT HELIX 368 371
FT TURN 372 372

```

Query Match 100.0%; Score 40; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.73;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7

Db 162 SSODYFH 168

```

RESULT 3.
AMYB_ASPAW STANDARD; PRT: 499 AA.
ID AMYB_ASPAW
AC 002906;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alpha-amylase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase B).
GN AMYB.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DVK143F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
RT from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212(1990).
CC -I CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -I COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -I SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X52756; CAA36967.1; -.
CC DR HSSP; P10529; 7YAA.
CC DR InterPro; IPR000461; Alpha-amylase.
CC DR Pfam; PF00128; alpha-amylase.1.
CC DR Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
CC Glycoprotein; Signal; Multigene family.
CC KW SIGNAL 1 21
CC FT CHAIN 22 499 ALPHA-AMYLASE B.
CC FT ACT_SITE 227 227 BY SIMILARITY.
CC FT ACT_SITE 231 231 BY SIMILARITY.
CC FT ACT_SITE 251 251 BY SIMILARITY.
CC FT ACT_SITE 318 318 BY SIMILARITY.
CC FT DISULFID 51 59 BY SIMILARITY.
CC FT DISULFID 171 185 BY SIMILARITY.
CC FT DISULFID 261 304 BY SIMILARITY.
CC FT DISULFID 461 496 BY SIMILARITY.
CC FT CARBOHYD 218 218 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 499 AA; 54921 MW; 740B96B11BC01A8A CRC64;

```

Query Match 100.0%; Score 40; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.73;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7

Db 162 SSODYFH 168

RESULT 4
 AMY_ASPSH

```

ID AMY_ASBSH STANDARD; PRT; 499 AA.
AC P30292;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
GN AMY.
OS Aspergillus shirousami.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eukaryotes; Trichocommata; mitosporic Trichocommata; Aspergillus.
OX NCBI_TaxID=5070;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92323146; PubMed=1368777;
RA Shibuya I., Tamura G., Ishikawa T., Hara S.;
RT "Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its
RT expression in Saccharomyces cerevisiae."
RL Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D10461; BAA01255.1; -.
DR PIR: J50663; J50663.
DR HSSP: P10529; J50663.
DR InterPro: IPR000461; Alpha_amylase.
DR Pfam: PF00128; alpha-amylase.1.
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 499 ALPHA-AMYLASE.
FT ACT_SITE 227 227
FT ACT_SITE 231 231 BY SIMILARITY.
FT ACT_SITE 251 251 BY SIMILARITY.
FT ACT_SITE 318 318 BY SIMILARITY.
FT DISULFID 51 59 BY SIMILARITY.
FT DISULFID 171 185 BY SIMILARITY.
FT DISULFID 261 304 BY SIMILARITY.
FT DISULFID 461 496 BY SIMILARITY.
FT CARBOHYD 218 218 N-LINKED (GLCNAC... ) (BY SIMILARITY).
SQ SEQUENCE 499 AA; 54852 MW; 1FB7AE50DA01C03F CRC64;

Query Match 100.0%; Score 40; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
DB 162 SSODYFH 168

```

```

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=57278;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J1;
RA Nicholas J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA
CC REPLICATION.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U43400; AAC54703.1; -.
DR InterPro: IPR000635; Viral_DNA_bind.
DR Pfam: PF00747; Viral_DNA_bp.1.
KW DNA-binding; DNA replication; Zinc-finger; Nuclear protein.
FT ZN_FING 459 475 C2HC-TYPE.
SQ SEQUENCE 1131 AA; 129008 MW; 52C97388D1B6D04F CRC64;

Query Match 82.5%; Score 33; DB 1; Length 1131;
Best Local Similarity 71.4%; Pred. No. 48;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
DB 943 ASODYFH 949

RESULT 6
AC AGH_ARMVU STANDARD; PRT; 144 AA.
ID AGH_ARMVU
AC Q908R2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Androgenic gland hormone precursor.
DE Armadillidium vulgare (Woodlice) (Pillbugs).
OS Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Peracarida; Isopoda; Oniscidea; Armadillidiidae;
OC Armadillidium.
OX NCBI_TaxID=13347;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Androgenic gland;
RC MEDLINE=20001935; PubMed=10529379;
RX Okuno A., Hasegawa Y., Ohira T., Katakura Y., Nagasawa H.;
RA "Characterization and cDNA cloning of androgenic gland hormone of the
RA terrestrial isopod Armadillidium vulgare."
RL Biochem. Biophys. Res. Commun. 264:419-423(1999).
RN [2]
RP PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE.
RC TISSUE=Androgenic gland;
RX MEDLINE=99339941; PubMed=10411634;
RA Martin G., Sorokine O., Moniatte M., Bulet P., Hetru C.,
RA Van Dorsselaer A.;
RT "The structure of a glycosylated protein hormone responsible for sex
RT determination in the isopod, Armadillidium vulgare."
RL Eur. J. Biochem. 262:727-736(1999).
CC -1- FUNCTION: CONTROLS SEX DIFFERENTIATION AND THE FORMATION OF MALE
CC APPENDAGES, SPERMATOGENESIS, PIGMENTATION, AND MALE SPECIFIC
CC BEHAVIOR.
CC -1- SUBCELLULAR LOCATION: Secreted.

```

```

CC -1- TISSUE SPECIFICITY: ANDROGENIC GLAND.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: AB029615; BAA6893.1; -.
DR Glycosylated; Q9U8R2; -.
KW Signal; Glycoprotein; Hormone; Sexual differentiation;
KW Spermatogenesis.
FT SIGNAL 1 21
FT CHAIN 22 65 ANDROGENIC GLAND HORMONE B CHAIN.
FT PROPEP 68 113 C PEPTIDE.
FT CHAIN 116 144 ANDROGENIC GLAND HORMONE A CHAIN.
FT DISULFID 33 123 INTERCHAIN (OR 124).
FT DISULFID 42 59 OR 44.
FT DISULFID 44 141 INTERCHAIN (OR 42).
FT DISULFID 124 132 OR 123.
FT CARBOHYD 133 133 N-LINKED (GLCNAC. .) (COMPLEX).
FT /FTID-CAR.000163.
SQ SEQUENCE 144 AA; 16893 MW; ED1541134628500B CRC64;

```

```

Query Match 80.0%; Score 32; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 ODYFH 7
    |||||
DB 82 ODYFH 86

```

```

RESULT 7
HXD8_MOUSE STANDARD; PRT; 289 AA.
AC P23453;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Homeobox protein Hox-D8 (Hox-4.3) .
OS HOXD8 OR HOXD-8 OR HOX-4.3.
CN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91209232; PubMed=1982431;
RX Izpisua-Belmonte J.-C., Dolle P., Renucci A., Zappavigna V.,
RX Falkenstein H., Duboule D.;
RA "Primary structure and embryonic expression pattern of the mouse
RT Hox-4.3 homeobox gene.";
RL Development 110:733-745(1990).
RN [2]
RP SEQUENCE OF 191-289 FROM N.A.
RX MEDLINE=91274361; PubMed=1675873;
RX Sadoul R., Featherstone M.;
RA "Sequence analysis of the homeobox-containing exon of the murine
RT Hox-4.3 homeobox gene.";
RL Biochim. Biophys. Acta 1089:259-261(1991).
RN [3]
RP SEQUENCE OF 195-254 FROM N.A.
RX MEDLINE=92073356; PubMed=1683707;
RX Singh G., Kaur S., Stock J.L., Jenkins N.A., Gilbert D.J.,
RX Copeland N.G., Potter S.S.;
RA "Identification of 10 murine homeobox genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10706-10710(1991).
RN [4]
RP SEQUENCE OF 192-260 FROM N.A.

```

```

RX MEDLINE=92212934; PubMed=1348361;
RA Nazareali A., Kim Y., Nirenberg M.;
RT "Hox-1.11 and Hox-4.9 homeobox genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2883-2887(1992).
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -1- SIMILARITY: BELONGS TO THE ANTP FAMILY OF HOMEOBOX PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: X56561; CA39911.1; -.
DR EMBL: M87803; AAA37852.1; -.
DR PIR: A38810; A38810.
DR PIR: A41605; A41605.
DR PIR: A43562; A43562.
DR PIR: S16177; S16177.
DR PIR: S15521; S15521.
DR HSSP: P02833; 9ANT.
DR TRANSFAC: T01426; -.
DR MGD; MGI:96209; Hoxd8.
DR InterPro: IPR001827; Antennapedia.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00025; ANTENNAPEDIA.
DR PRINTS: PR00024; HOMEOBOX.
DR SMART: SM00389; HOX.1.
DR PROSITE: PS00027; HOMEOBOX_1.1.
DR PROSITE: PS00032; ANTENNAPEDIA.1.
DR PROSITE: PS50071; HOMEOBOX_2.1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 15 28 POLY-ALA.
FT DOMAIN 62 89 GLY/PRO-RICH.
FT DOMAIN 108 117 POLY-PRO.
FT DOMAIN 183 188 ANTP-TYPE HEXAPEPTIDE.
FT DNA_BIND 195 254 HOMEOBOX.
FT CONFLICT 207 208 TL -> RV (IN REF. 1).
FT CONFLICT 231 231 T -> S (IN REF. 1).
FT CONFLICT 265 266 EA -> DG (IN REF. 1).
FT CONFLICT 275 275 A -> V (IN REF. 2).
SQ SEQUENCE 289 AA; 31410 MW; 5783099B952BDFE CRC64;

```

```

Query Match 80.0%; Score 32; DB 1; Length 289;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 ODYFH 7
    |||||
DB 90 ODYFH 94

```

```

RESULT 8
VGLM_HSVBC STANDARD; PRT; 411 AA.
ID VGLM_HSVBC
AC P52370;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Glycoprotein M.
GN GM OR U110.
OS Bovine herpesvirus type 1 (strain Cooper).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphanterpesvirinae; Varicelloviruses.
OX NCBI_TaxID=10323;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9531343; PubMed=7793062;
RA Vilek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
RA Letchworth G.J., Schwyzer M.;
RT "Nucleotide sequence analysis of a 30-kb region of the bovine
RT herpesvirus 1 genome which exhibits a colinear gene arrangement with
RT the UL21 to UL4 genes of herpes simplex virus.";
RL Virology 210:100-108(1995).
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN M.
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN M.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z48053; CAAB8123.1;
DR InterPro: IPR000785; Herpes_glycop.
DR Pfam: PF01528; Herpes_glycop.1.
DR PRINTS: PR00333; HSVINTEGRIMP.
DR Transmembrane; Glycoprotein.
FT TRANSMEM 14 34 POTENTIAL.
FT TRANSMEM 89 109 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 159 179 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 249 269 POTENTIAL.
FT TRANSMEM 277 297 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT CARBOHYD 57 57 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 411 AA: 43629 MW: 20F156DA9FA0158C CRC64;

Query Match 80.0%; Score 32; DB 1; Length 411;
Best Local Similarity 71.4%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSODYPH 7
DB 295 ASHDFH 301

RESULT 9
AKH1_ECOLI STANDARD: PRT; 820 AA.
AC P00561; Q47659;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bifunctional aspartokinase/homoserine dehydrogenase I (AKI-HDI)
DE [Includes: Aspartokinase I (EC 2.7.2.4); Homoserine dehydrogenase I
DE (EC 1.1.1.3)].
GN THRA OR THRA1 OR THRA2 OR B0002.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
ON [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81077247; PubMed=7003595;
RA Katinka M., Cossart P., Sibilli L., Saint-Girons I., Chalvignac M.A.,
RA le Bras G., Cohen G.N., Yaniv M.;
RT "Nucleotide sequence of the thra gene of Escherichia coli.";
RT Proc. Natl. Acad. Sci. U.S.A. 77:5730-5733(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-K12;
RX MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,

```

```

RA Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region.";
RL Nucleic Acids Res. 20:3305-3308(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=80077291; PubMed=390305;
RA Cossart P., Katinka M., Yaniv M.;
RT "Construction and expression of a hybrid plasmid containing the
RT Escherichia coli thra and thrB genes.";
RL Mol. Gen. Genet. 175:39-44(1979).
RN [5]
RP SEQUENCE OF 51-129.
RX MEDLINE=80043179; PubMed=387092;
RA Sibilli L., le Bras G., Cossart P., Chalvignac M.A., le Bras G.,
RA Briley P.A., Cohen G.N.;
RT "The primary structure of Escherichia coli K 12 aspartokinase I-
RT homoserine dehydrogenase I: sequence of cyanogen bromide peptide CB
RT 3.".
RL Biochimie 61:733-739(1979).
RN [6]
RP REVISION TO 11.
RX MEDLINE=8135751; PubMed=6298218;
RA Cossart P.;
RT Unpublished results, cited by:
RL Zakin M.M., Duchange N., Ferraz P., Cohen G.N.;
RL J. Biol. Chem. 258:3028-3031(1983).
CC -1- CATALYTIC ACTIVITY: L-homoserine + NAD(P)+ = L-aspartate 4-
CC semialdehyde + NAD(P)H.
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate = ADP + 4-phospho-L-
CC aspartate.
CC -1- ENZYME REGULATION: THE ENZYME ACTIVITIES ARE REGULATED
CC ALLOSTERICALLY BY L-THREONINE.
CC -1- PATHWAY: CATALYZES 2 NONCONSECUTIVE REACTIONS IN THE COMMON
CC BIOSYNTHETIC PATHWAY LEADING FROM ASP TO DIAMINOPIMELATE AND LYS,
CC TO MET, AND TO THR AND ILE.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- MISCELLANEOUS: ASPARTOKINASE II-HOMOSERINE DEHYDROGENASE II AND
CC ASPARTOKINASE III ALSO CATALYZE THE SAME REACTION(S).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE
CC ASPARTOKINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE
CC HOMOSERINE DEHYDROGENASE FAMILY.
CC -----
DR EMBL: V00361; CAA23660.1; ALT_SEQ.
DR EMBL: J01706; AAA83914.1;
DR EMBL: X68872; CAA8734.1;
DR EMBL: D10483; BAA01286.1;
DR EMBL: U14003; AAA97301.1;
DR EMBL: AE000111; AAC73113.1;
DR EMBL: V00360; CAA23659.1;
DR EMBL: M28570; AAA24673.1;
DR EMBL: M10644; AAA24671.1;
DR PIR: A00671; DECK.
DR PIR: S40531; S40531.
DR Ecogene; EG10998; thra.

```

```

DR InterPro; IPR002912; ACT.
DR InterPro; IPR001048; Aakkinase.
DR InterPro; IPR001341; Aspartokinase.
DR InterPro; IPR001342; Homoserine_dh.
DR Pfam; PF00696; aakkinase; 1.
DR Pfam; PF01842; ACT; 2.
DR Pfam; PF00742; Homoserine_dh; 1.
DR PROSITE; PS00324; ASPARTOKINASE; 1.
DR PROSITE; PS01042; HOMOSERINE DEHYDROGENASE; 1.
DR TRANSFERASE; Kinase; Oxidoreductase; Threonine biosynthesis; NADP;
KW Allosteric enzyme; Multifunctional enzyme; Complete proteome.
FT DOMAIN 1 249 ASPARTOKINASE I.
FT DOMAIN 250 470 INTERFERE.
FT NP_BIND 471 820 HOMOSERINE DEHYDROGENASE I.
FT NP_BIND 471 478 NADP (POTENTIAL).
FT CONFLICT 113 113 Q -> E (IN REF. 5).
FT CONFLICT 230 230 N -> D (IN REF. 3).
FT CONFLICT 375 375 L -> Q (IN REF. 3).
FT CONFLICT 393 393 A -> T (IN REF. 3).
FT CONFLICT 406 406 L -> M (IN REF. 3).
FT CONFLICT 553 553 N -> D (IN REF. 3 AND 4).
FT CONFLICT 587 588 DY -> IT (IN REF. 4).
FT CONFLICT 607 607 I -> T (IN REF. 3).
FT CONFLICT 658 658 R -> T (IN REF. 3).
SQ SEQUENCE 820 AA: 89122 MW: 84DE917A5703F6E7 CRC64:

Query Match      80.0%; Score 32; DB 1; Length 820;
Best Local Similarity 71.4%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSODYFH 7
DB 584 SSMDDYH 590

RESULT 10
ID UBP8_HUMAN STANDARD; PRT; 1118 AA.
AC P40818;
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase 8 (EC 3.1.2.15) (Ubiquitin
DE thiolesterase 8) (Ubiquitin-specific processing protease 8)
DE (Deubiquitinating enzyme 8).
GN USP8 OR KIAA0055.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=bone marrow;
RX MEDLINE=96051398; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima K.-i., Tanaka A., Sato S.,
RA Seki N., Kawarabayashi Y., Ishikawa K.-i., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (K1A0041-K1A0080) deduced by
RT analysis of cDNA clones from human cell line KG-1."
RL DNA Res. 1:223-229(1994).
RL -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
RL ubiquitin + a thiol.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19; ALSO KNOWN AS
CC FAMILY 2 OF UBIQUITIN CARBOXYL-TERMINAL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; D29956; BAA06225.1; -.
DR MEROPS; C19.011; -.
DR MIM; 603158; -.
DR InterPro; IPR001763; Rhodanese_domain.
DR InterPro; IPR001394; UCH-2.
DR Pfam; PF00581; Rhodanese; 1.
DR Pfam; PF00442; UCH-1; 1.
DR Pfam; PF00442; UCH-2; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS02335; UCH_2_3; 1.
KW Ubiquitin conjugation; Hydrolase; Thiol protease; Multigene family.
FT ACT_SITE 786 786 BY SIMILARITY.
FT ACT_SITE 1059 1059 BY SIMILARITY.
FT ACT_SITE 1067 1067 BY SIMILARITY.
SQ SEQUENCE 1118 AA: 127523 MW: 8B84B7A842P9A9A CRC64:

Query Match      80.0%; Score 32; DB 1; Length 1118;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QDYFH 7
DB 83 QDYFH 87

RESULT 11
ID YMF8_YEAST STANDARD; PRT; 104 AA.
AC Q04964;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 11.8 kDa protein in OG1-CNA2 intergenic region.
DE YML058W OR YMF958.04.
GN Saccharomyces cerevisiae (Baker's yeast).
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Devlin K., Churcher C., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46729; CAA86717.1; -.
DR SGD; S0004523; YML058W.
KW Hypothetical protein.
SQ SEQUENCE 104 AA: 11834 MW: 02735C9178F431B2 CRC64:

Query Match      77.5%; Score 31; DB 1; Length 104;
Best Local Similarity 71.4%; Pred. No. 8;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
DB 3 SSODYFH 9

RESULT 12
CCG7_HUMAN

```

```

ID  CCG7 HUMAN          STANDARD;          PRT;          198 AA.
AC  O9BX1;
DT  01-MAR-2002 (Rel. 41, Created)
DT  01-MAR-2002 (Rel. 41, Last sequence update)
DT  01-MAR-2002 (Rel. 41, Last annotation update)
DE  Voltage-dependent calcium channel gamma-7 subunit (Neuronal voltage-
DE  gated calcium channel gamma-7 subunit).
GN  CACNG7.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX  NCBI-TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=21100909; PubMed=11170751;
RX  Burgess D.L., Gefrides L.A., Foreman P.J., Noebels J.L.;
RT  "A cluster of three novel Ca(2+) channel gamma subunit genes on
RT  chromosome 19q13.4: evolution and expression profile of the gamma
RT  subunit gene family.";
RL  Genomics 71:339-350(2001).
CC  -1- FUNCTION: THOUGHT TO STABILIZE THE CALCIUM CHANNEL IN AN
CC  INACTIVATED (CLOSED) STATE (BY SIMILARITY).
CC  -1- SUBUNIT: THE L-TYPE CALCIUM CHANNEL IS COMPOSED OF FIVE SUBUNITS:
CC  ALPHA-1, ALPHA-2/DELTA, BETA AND GAMMA.
CC  -1- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).
CC  -1- SIMILARITY: BELONGS TO THE PMP-22 / EMP / MP20 FAMILY. CACNG
CC  SUPERFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF288387; AAK20030.1; -.
DR  InterPro; IPR004031; PMP22_Claudin.
DR  Pfam; PF00822; PMP22_Claudin; 1.
KW  Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW  Calcium channel.
FT  TRANSMEM 8 28 POTENTIAL.
FT  TRANSMEM 129 149 POTENTIAL.
FT  SEQUENCE 198 AA; 22041 MW; AF755430E0A0073B CRC64;
SQ
Query Match 77.5%; Score 31; DB 1; Length 198;
Best local similarity 71.4%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 SSODTFH 7
DB 166 SSEOYFH 172
ID 11: 111
ID 166 SSEOYFH 172
RESULT 13
LBM_DROME
ID LBM_DROME STANDARD; PRT; 208 AA.
AC Q24188; Q9Y4H0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Late bloomer protein.
GN LBM OR LBL OR CG2374.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI-TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96175586; PubMed=8596956;

```

```

RA Kopczyński C.C., Davis G.W., Goodman C.S.;
RT "A neuronal tetraspanin, encoded by late bloomer, that facilitates
RT synapse formation."
RL Science 271:1867-1870(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RX Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RX Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RX George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RX Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RX Brandon R.C., Rogers Y.-H.C., Blazer V., Holt R.A., Davis R.,
RX Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RX Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RX Beeson K.V., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
RX Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
RX Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RX Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RX de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RX Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RX Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RX Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RX Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RX Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RX Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RX Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
RX Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RX Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RX Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RX Merkllov G., Milshina N.V., Mobarry C., Morris J., Mostrel A.,
RX Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
RX Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RX Palazolo M., Pittman K.S., Pan S., Pollard J., Pui Y., Reese M.G.,
RX Raebert K., Remington I., Saunders R.D.C., Scheeler F., Shen H.,
RX Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RX Spierkas R., Spradling A.C., Stapleton M., Strong R., Sun E.,
RX Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RX Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RX Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RX Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RX Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RX Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: FACILITATES SYNAPSE FORMATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO THE TETRASPANIN (TM4SF) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U49081; AAA98512.1; -.
DR EMBL; AEO03842; AAF59302.1; -.
DR FlyBase; FBgn016032; Lbm.
DR InterPro; IPR000301; Transmem_4.
DR Pfam; PF00335; transmembrane4; 1.
KW Transmembrane.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 61 61 POTENTIAL.
FT TRANSMEM 67 87 POTENTIAL.
FT TRANSMEM 174 194 POTENTIAL.
SQ SEQUENCE 208 AA; 23005 MW; 5EC36525287B7AC9 CRC64;
Query Match 77.5%; Score 31; DB 1; Length 208;
Best local similarity 71.4%; Pred. No. 19;

```

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSODYFH 7
1:1111
Db 128 SAODYFH 134

RESULT 14

TIAL_MOUSE

ID TIAL_MOUSE STANDARD; PRT; 386 AA.

AC P52912;

DT 01-OCT-1996 (rel. 34, Created)

DT 01-OCT-1996 (rel. 34, Last annotation update)

DE 16-OCT-2001 (rel. 40, Last annotation update)

DE Nucleolysin TIA-1 (RNA-binding protein TIA-1).

GN TIA1 OR TIA.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

RN 11

RP SEQUENCE FROM N.A.

RC TISSUE=Blood;

RA MEDLINE=96322730; PubMed=8759725;

RA Lowin B., French L., Martinou J.C., Tschoopp J.

RT "Expression of the CTL-associated protein TIA-1 during murine

embryogenesis.";

RL J. Immunol. 157:1448-1454(1996).

RN 12

RP SEQUENCE FROM N.A.

RA MEDLINE=97025375; PubMed=8871565;

RA Beck A.R.P., Medley O.G., O'Brien S., Anderson P., Streuli M.;

RT "Structure, tissue distribution and genomic organization of the

murine RRM-type RNA binding proteins TIA-1 and TIAR.";

RL Nucleic Acids Res. 24:3829-3836(1996).

CC -1- FUNCTION: RNA-BINDING PROTEIN. POSSESSES NUCLEOLYTIC ACTIVITY

AGAINST CYTOTOXIC LYMPHOCYTE TARGET CELLS. MAY BE INVOLVED IN

CC APOPTOSIS (BY SIMILARITY).

CC SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES OF CYTOLYTIC

CC T-LYMPHOCYTES (BY SIMILARITY).

CC -1- SIMILARITY: CONTAINS 3 RNA RECOGNITION MOTIFS (RRM).

CC -1- SIMILARITY: TO NUCLEAR RNA-BINDING PROTEINS. STRONGEST TO TIAR.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC

CC EMBL; U00689; AAA03711.1; -

DR EMBL; U55867; AAC52871.1; -

DR MGD; MGI:107914; Tial.

DR InterPro; IPR000504; RRM.

DR Pfam; PF00076; RRM; 3.

DR SMART; SM00360; RRM; 3.

DR PROSITE; PS0102; RRM; 3.

DR PROSITE; PS00030; RRM_RNP_1; 2.

DR RNA-binding; Apoptosis; Repeat.

FT DOMAIN 7 83 RNA-BINDING (RRM) 1.

FT DOMAIN 106 184 RNA-BINDING (RRM) 2.

FT DOMAIN 214 286 RNA-BINDING (RRM) 3.

SO SEQUENCE 386 AA; 42800 MW; 51BFPE9DE046D3AE CRC64;

Query Match 77.5%; Score 31; DB 1; Length 386;

Best Local Similarity 83.3%; Pred. No. 37;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSODYFH 7
1111111

Db 102 SSODYFH 107

RESULT 15

DDX5_HUMAN

ID DDX5_HUMAN STANDARD; PRT; 614 AA.

AC P17844;

DT 01-AUG-1990 (rel. 15, Created)

DT 01-AUG-1990 (rel. 15, Last sequence update)

DE 16-OCT-2001 (rel. 40, Last annotation update)

DE Probable RNA-dependent helicase p68 (DEAD-box protein p68) (DEAD-box

protein 5).

GN DDX5 OR HLR1 OR HELR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxId=9606;

RN 11

RP SEQUENCE FROM N.A.

RA MEDLINE=90272412; PubMed=2349099;

RA Hloch P., Stahl H.;

RT "Complete cDNA sequence of the human p68 protein.";

RL Nucleic Acids Res. 18:3045-3045(1990).

RN 12

RP SEQUENCE FROM N.A.

RA MEDLINE=91141480; PubMed=1996094;

RA 1990 R.D., Jamieson D.D., McNeill S.A., Southgate J., McPheat J.,

lane D.P.;

RT "p68 RNA helicase: identification of a nucleolar form and cloning of

RT related genes containing a conserved intron in yeasts.";

RL Mol. Cell. Biol. 11:1326-1333(1991).

RN 13

RP SEQUENCE FROM N.A.

RA Hloch P., Roessler O.G., Weitzenecker T., Stahl H.;

RL Submitted (Nov-1997) to the EMBL/GenBank/DBJ databases.

RN 14

RP SEQUENCE OF 21-614 FROM N.A.

RA MEDLINE=88189326; PubMed=2451786;

RA Ford M.J., Anton I.A., Lane D.P.;

RT "Nuclear protein with sequence homology to translation initiation

RT factor eIF-4A.";

RL Nature 332:736-738(1988).

CC -1- FUNCTION: RNA-DEPENDENT ATPASE ACTIVITY. THE RATE OF ATP

CC HYDROLYSIS IS HIGHLY STIMULATED BY SINGLE-STRANDED RNA.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX5/DDX17

CC SUBFAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC

CC EMBL; X52104; CAA36324.1; -

DR EMBL; AF015812; AAB84094.1; -

DR EMBL; X15729; CAA33751.1; -

DR PIR; S06377; S06377.

DR PIR; S10181; S10181.

DR PIR; S14045; S14045.

DR Harbus/Chent-ZDPAE; 1605; NEPHGE.

DR HSP; Q58083; IHV8.

DR MIM; 180630; -

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR000629; DEAD_ATP_helicase.

DR InterPro; IPR001650; Helicase_C.

DR Pfam; PF00270; DEAD; 1.

DR Pfam; pf00271; helicase_C; 1.

DR SMART; SM00487; DEXDC; 1.

DR SMART; SM00490; HELIC; 1.

DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.

DR ATP-binding; RNA-binding; Helicase; Nuclear protein.

NP BIND 138 145 ATP (BY SIMILARITY).
 SITE 248 251 DEAD BOX
 SQ SEQUENCE 614 AA; 69148 MW; 84DF684FD6871594 CRC64;

Query Match 77.5%; Score 31; DB 1; Length 614;
 Best Local Similarity 71.4%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSQDYEH 7
 11:111
 Db 421 SSQDYEH 427

Search completed: June 13, 2002, 09:11:53
 Job time: 1033 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:10 ; Search time 125.81 Seconds
(without alignments)
9.625 Million cell updates/sec

Title: US-09-710-339-2_COPY_161_167

Perfect score: 40
Sequence: 1 SSQDYFH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._virus:*
16: sp._bacteriophage:*
17: sp._archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	40	100.0	499	3	Q96TH4	Q96TH4 aspergillus
2	40	100.0	624	3	Q01117	Q01117 lipomyces k
3	35	87.5	760	5	Q9XJ0	Q9XJ0 giardia lam
4	34	85.0	130	10	Q23373	Q23373 arabidopsis
5	33	82.5	420	12	O55525	O55525 equine herp
6	33	82.5	420	12	O39309	O39309 equine herp
7	33	82.5	585	11	Q9D718	Q9D718 mus musculi
8	33	82.5	601	5	O17552	O17552 caenorhabdi
9	33	82.5	1131	12	O56282	O56282 human sapie
10	33	82.5	1731	4	Q9P230	Q9P230 homo sapien
11	33	82.5	3124	4	Q96L91	Q96L91 homo sapien
12	33	82.5	4351	11	O88277	O88277 rattus norv
13	32	80.0	163	11	Q9C0Q8	Q9C0Q8 mus musculu
14	32	80.0	180	16	Q97SK2	Q97SK2 streptococc
15	32	80.0	215	14	P97975	P97975 unidentified
16	32	80.0	218	11	Q9CUN8	Q9CUN8 mus musculu

17	32	80.0	317	16	Q9K764	Q9K764 bacillus ha
18	32	80.0	351	10	Q23041	Q23041 arabidopsis
19	32	80.0	419	5	O01765	O01765 caenorhabdi
20	32	80.0	438	12	O39493	O39493 bovine herp
21	32	80.0	482	17	O97ZF2	O97ZF2 sulfolobus
22	32	80.0	529	5	Q9YK03	Q9YK03 drosophila
23	32	80.0	575	11	Q31ZG4	Q31ZG4 mus musculu
24	32	80.0	583	4	Q92802	Q92802 homo sapien
25	32	80.0	685	3	Q08683	Q08683 saccharomyc
26	31	77.5	67	12	Q84487	Q84487 paramecium
27	31	77.5	172	11	Q90XU1	Q90XU1 mesocricetu
28	31	77.5	172	11	Q99MG7	Q99MG7 mesocricetu
29	31	77.5	200	6	Q28465	Q28465 macropus eu
30	31	77.5	214	4	Q96B58	Q96B58 homo sapien
31	31	77.5	214	16	Q9RX20	Q9RX20 delnoccocus
32	31	77.5	219	6	Q95LS2	Q95LS2 macaca fasc
33	31	77.5	261	16	Q92AE6	Q92AE6 listeria in
34	31	77.5	304	11	Q9CSD9	Q9CSD9 mus musculu
35	31	77.5	388	13	Q90WNO	Q90WNO xenopus lae
36	31	77.5	461	16	Q92750	Q92750 chlamydia p
37	31	77.5	490	3	Q9UY07	Q9UY07 emericella
38	31	77.5	595	13	Q9W744	Q9W744 gallus gall
39	31	77.5	597	2	Q952K2	Q952K2 streptomyce
40	31	77.5	608	13	Q918S3	Q918S3 xenopus lae
41	31	77.5	668	10	Q9LDM7	Q9LDM7 arabidopsis
42	31	77.5	761	5	Q9NL40	Q9NL40 asterina pe
43	31	77.5	825	6	Q28477	Q28477 macaca fasc
44	31	77.5	858	6	O19056	O19056 papio anudi
45	31	77.5	893	5	O16221	O16221 caenorhabdi

ALIGNMENTS

RESULT 1
ID Q96TH4 PRELIMINARY; PRT; 499 AA.
AC Q96TH4;
DT 01-DEC-2001 (TREMREL. 19, Created)
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE TAKA-AMYLASE A (EC 3.2.1.1) (ALPHA-AMYLASE).
GN AMYA OR AMY1.
OS Aspergillus oryzae, and
OS Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
CX NCBI_Taxid=5062, 5059;
LN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=A.oryzae; STRAIN=R1840;
RX MEDLINE=2028310; PubMed=10830498;
RA Gomi K., Akeno T., Mineokki T., Ozeki K., Kumagai C., Okazaki N.,
Iimura Y.,
RT "Molecular cloning and characterization of a transcriptional activator
RT gene, amyR, involved in the amyloid gene expression in Aspergillus
RT oryzae.";
RL Biosci. Biotechnol. Biochem. 64:816-827(2000).
LN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=A.flavus; STRAIN=86-10D;
RA Fakhoury A.K., Woloshuk C.P.,
RT "Amy1, the alpha-amylase gene of Aspergillus flavus: Involvement in
RT aflatoxin biosynthesis in maize kernels.";
RL Phytopathology 89:908-914(1999).
DR EMBL; AB021876; BAA95703.1; -;
DR EMBL; AF139925; AAF14264.1; -;
SQ SEQUENCE 499 AA; 54804 MW; EEFA2ADA71D20DA9 CRC64;

Query Match 100.0%; Score 40; DB 3; Length 499;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 SSODYFH 7
Db      162 SSODYFH 168

RESULT  2
ID      001117      PRELIMINARY;      PRT;      624 AA.
AC      001117;
DT      01-NOV-1998 (TREMblrel. 08, Created)
DT      01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT      01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE      ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1)
DE      (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE) (LKA1).
GN      LKA1.
OS      Lipomyces kononenkoae.
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Lipomycetaceae; Lipomyces.
OX      NCBI_TaxID=34357;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=IGC4052B;
RX      MEDLINE=96105202; PubMed=8529895;
RA      Steyn A.J.C., Marmur J., Pretorius I.S.;
RT      "Cloning, sequence analysis and expression in yeasts of a cDNA
RL      containing a Lipomyces kononenkoae alpha-amylase-encoding gene.";
RN      Gene 166:65-71(1995).
RN      [2]
RP      SEQUENCE OF 29-44.
RC      STRAIN=IGC4052B; PubMed=8593683;
RX      MEDLINE=96133108;
RA      Steyn A.J.C., Pretorius I.S.;
RT      "Characterization of a novel alpha-amylase from Lipomyces kononenkoae
RT      and expression of its gene (LKA1) in Saccharomyces cerevisiae.";
RN      Curr. Genet. 28:526-533(1995).
CC      -1- FUNCTION: LIBERATES REDUCING GROUPS FROM POLYMERS CONTAINING BOTH
CC      1,4-ALPHA AND 1,6-ALPHA BONDS.
CC      -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC      LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC      -1- CORFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
CC      SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC      -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES. ALSO
CC      KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR      EMBL: U30376; AAC49622.1; ALT_INIT.
DR      HSPF: P10529; 77NA.
DR      InterPro: IPR00461; Alpha-amylase.
DR      Pfam: PF00128; alpha-amylase; 1.
KW      Hydrolase; Glycosidase; Carbohydrate metabolism; Signal; Calcium;
KW      glycoprotein.
FT      SIGNAL      1      28
FT      CHAIN      29      624      ALPHA-AMYLASE.
FT      ACT_SITE    353      353      BY SIMILARITY.
FT      ACT_SITE    357      357      BY SIMILARITY.
FT      ACT_SITE    377      377      BY SIMILARITY.
FT      ACT_SITE    444      444      BY SIMILARITY.
FT      ACT_SITE    444      444      BY SIMILARITY.
FT      DISULFID    177      185      BY SIMILARITY.
FT      DISULFID    297      311      BY SIMILARITY.
FT      DISULFID    387      430      BY SIMILARITY.
FT      DISULFID    587      622      BY SIMILARITY.
FT      DISULFID    304      344      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD    344      344      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD    344      344      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE      624 AA; 68876 MW; 87EB16534F5A9A9F CRC64;

Query Match      100.0%; Score 40; DB 3; Length 624;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSODYFH 7
Db      288 SSODYFH 294

```

```

RESULT  3
ID      09XZJ0      PRELIMINARY;      PRT;      760 AA.
AC      09XZJ0;
DT      01-NOV-1999 (TREMblrel. 12, Created)
DT      01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT      01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE      ANKYRIN-LIKE PROTEIN.
DE      Giardia lamblia (Giardia intestinalis).
OS      Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.
OX      NCBI_TaxID=5741;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=PORTLAND;
RA      Sekyere E.O., Edwards M.R., Stewart T.S., Schofield P.J., Knodler L.;
RT      "Ankyrin-like protein from Giardia intestinalis.";
RL      Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF144322; AAD28486.1; -.
DR      InterPro: IPR002110; ANK.
DR      Pfam: PR00023; ank; 2.
DR      SMART: SM00248; ANK; 1.
KW      ANK repeat; Repeat.
SQ      SEQUENCE      760 AA; 85542 MW; 12E54841019CD475 CRC64;

Query Match      87.5%; Score 35; DB 5; Length 760;
Best Local Similarity 85.7%; Pred. No. 45;
Matches      6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SSODYFH 7
Db      5 SSODYFH 11

```

```

RESULT  4
ID      023373      PRELIMINARY;      PRT;      130 AA.
AC      023373;
DT      01-JAN-1998 (TREMblrel. 05, Created)
DT      01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT      01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE      HYPOTHETICAL PROTEIN.
DE      AT4G15190.
GN      Arabidopsis thaliana (Mouse-ear cress).
OS      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terryn N.,
RA      Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
RA      Pulidomenech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
RA      Jones J., Palme K., Ansoorge W., Delseny M., Bancroft I., Mewes H.W.,
RA      Schueller C., Chalwatzis N.,
RA      submitted (Jul-1997) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RA      EU Arabidopsis sequencing project;
RL      Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL: Z97338; CAB10298.1; -.
DR      EMBL: AL161540; CAB78561.1; -.
KW      Hypothetical protein.
SQ      SEQUENCE      130 AA; 14962 MW; DD87D2003030184 CRC64;

```

```

Query Match      85.0%; Score 34; DB 10; Length 130;
Best Local Similarity 85.7%; Pred. No. 11;
Matches      6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 SSODYFH 7

```

Db 38 SSADYFH 44

RESULT 5
ID 055525 PRELIMINARY; PRT; 420 AA.

AC 055525;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GLYCOPROTEIN I HOMOLOGUE.
OS Equine herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
NCBI_TaxID=10331;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TH20;
RX MEDLINE=98185635; PubMed=9524947;
RA Damiani A.M., Matsumura T., Yokoyama N., Maeda K., Miyazawa T.,
Kai C., Mikami T.;
RT "Nucleotide sequences of glycoprotein I and E genes of equine
herpesvirus type 4.";
RL J. Vet. Med. Sci. 60:219-225(1998).
DR EMBL: AB005749; BAA25021.1;
DR InterPro: IPR002874; Herpes-G1.
DR Pfam: PF01688; Herpes-G1;
SQ SEQUENCE 420 AA; 43696 MW; 1F64F5E179C2D916 CRC64;

Query Match 82.5%; Score 33; DB 12; Length 420;
Best Local Similarity 83.3%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SDDYFH 7
:|||||
Db 114 NDDYFH 119

RESULT 6
ID 039309 PRELIMINARY; PRT; 420 AA.
AC 039309;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE COUNTERPART OF HSV-1 GENE US7 AND VZV GENE 67.
GN 73.
OS Equine herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
NCBI_TaxID=10331;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NS80567;
RX MEDLINE=98264497; PubMed=9603335;
RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davidson A.J.;
RT "The DNA sequence of equine herpesvirus-4.";
RL J. Gen. Virol. 79:1197-1203(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NS80567;
RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davidson A.J.;
RT Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF030027; AAC59593.1;
DR InterPro: IPR002874; Herpes-G1.
DR Pfam: PF01688; Herpes-G1;
SQ SEQUENCE 420 AA; 45710 MW; 1EB944825DF3D852 CRC64;

Query Match 82.5%; Score 33; DB 12; Length 420;
Best Local Similarity 83.3%; Pred. No. 62;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SDDYFH 7
:|||||
Db 114 NDDYFH 119

RESULT 7
ID 09D718 PRELIMINARY; PRT; 585 AA.
AC 09D718;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 2310007D09R1K PROTEIN.
GN 2310007D09R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Schrml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Williams L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
Rayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK009199; BAB26134.1;
DR MGD: MGI:1919128; 2310007D09R1K.
DR InterPro: IPR001412; tRNA-synt.1.
DR PROSITE: PS00178; AA.TRNA-LIGASE.I; UNKNOWN.1.
SQ SEQUENCE 585 AA; 64324 MW; E3528EDAEB13FBFF CRC64;

Query Match 82.5%; Score 33; DB 11; Length 585;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SDDYFH 7
:|||||
Db 366 SDDYFH 372

RESULT 8
ID 017552 PRELIMINARY; PRT; 601 AA.
AC 017552;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BE10.3 PROTEIN.
GN BE10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nemata; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
NCBI_TaxID=6239;

RN [1]
 RP SEQUENCE FROM N.A.
 RA McMurray A.A.; 1997) to the EMBL/GenBank/DBJ databases.
 RL Submitted (Mar-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology";
 RL Science 283:2012-2018(1998).
 DR EMBL; 293372; CAB07545.1; -;
 SQ SEQUENCE 601 AA; 69490 MW; 30084EE8EC98B7E9 CRC64;

Query Match 82.5%; Score 33; DB 5; Length 601;
 Best Local Similarity 71.4%; Pred. No. 92;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 :|:|:|:|
 DB 453 NSRDYFH 459

RESULT 9
 ID 056282 PRELIMINARY; PRT; 1131 AA.
 AC 056282;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE SINGLE-STRANDED DNA-BINDING PROTEIN.
 GN U41.
 OS Human herpesvirus 7.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae;
 OX NCBT_TaxID=10372;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K;
 RA Megaw A.G., Rapaport D., Avidor B., Frenkel N., Davison A.J.;
 RX MEDLINE=9840941; PubMed=9581785;
 RT "The DNA sequence of the K strain of human herpesvirus 7.";
 RL Virology 244:119-132(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K;
 RA Megaw A.G., Frenkel N., Davison A.J.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF037218; AAC40755.1; -;
 DR Interpro; IPR000635; Viral_DNA_bind.
 DR Pfam; PF00747; Viral_DNA_dp; 1.
 KW DNA-binding.
 SQ SEQUENCE 1131 AA; 129018 MW; 165B3AAC703C82D CRC64;

Query Match 82.5%; Score 33; DB 12; Length 1131;
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
 :|:|:|:|
 DB 943 ASRDYFH 949

RESULT 10
 ID 09P230 PRELIMINARY; PRT; 1731 AA.
 AC 09P230;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE KIAA1498 PROTEIN (FRAGMENT).

GN KIAA1498.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBT_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20277482; PubMed=10819331;
 RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human
 genes. XVII. The complete sequences of 100 new cDNA clones from brain
 which code for large proteins in vitro";
 RL DNA Res. 7:143-150(2000).
 DR EMBL; AB040931; BAA96022.1; -;
 DR Interpro; IPR000330; SNEF2_N.
 DR Pfam; PF00176; SNEF2_N; 2.
 FT NON_TER 1
 FT NON_TER 1731
 SQ SEQUENCE 1731 AA; 188228 MW; D628E95F1021A8F CRC64;

Query Match 82.5%; Score 33; DB 4; Length 1731;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 SODYFH 7
 :|:|:|:|
 DB 1250 SODYFH 1255

RESULT 11
 ID 096L91 PRELIMINARY; PRT; 3124 AA.
 AC 096L91;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE P400 SWI2/SNF2-RELATED PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBT_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21400441; PubMed=11509179;
 RA Fuchs M., Gerber J., Drapkin R., Sif S., Ikura T., Ogrzyzko V.,
 RA Lane W.S., Nakatani Y., Livingston D.M.;
 RT "The p400 complex is an essential E1A transformation target.";
 RL Cell 106:297-307(2001).
 DR EMBL; AY044869; AAK97789.1; -;
 SQ SEQUENCE 3124 AA; 340146 MW; E8F57ED6C7BD01B9 CRC64;

Query Match 82.5%; Score 33; DB 4; Length 3124;
 Best Local Similarity 83.3%; Pred. No. 5.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 SODYFH 7
 :|:|:|:|
 DB 1249 SODYFH 1254

RESULT 12
 ID 088277 PRELIMINARY; PRT; 4351 AA.
 AC 088277;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MEGF1.
 GN MEGF1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX MEDLINE=98360089; PubMed=9693030;
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple EGF-
RL like motifs by motif-trap screening.";
RT Genomics 51:27-34(1998).
CC -1 SIMILARITY: CONTAINS 33 CADHERIN DOMAINS.
DR EMBL: AB011527; BAA32458.1; -.
DR HSSP: P08709; IBE9.
DR InterPro: IPR002126; Cadherin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001791; Laminin_G.
DR Pfam: PF00028; cadherin_33.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00054; Laminin_G_1.
DR PRINTS: PR00205; CADHERIN.
DR SMART: SM00112; CA_33.
DR SMART: SM00181; EGF_2.
DR SMART: SM00282; Lamg_1.
DR PROSITE: PS00232; CADHERIN_1; 1.
DR PROSITE: PS00268; CADHERIN_2; 33.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 2.
DR Calcium-binding; Cell adhesion; BGF-like domain; Glycoprotein.
SQ SEQUENCE 4351 AA; 480644 MW; C564E7F98EB388 CRC64;

Query Match          82.5%; Score 33; DB 11; Length 4351;
Best Local Similarity 71.4%; Pred. No. 7,9e+00;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SQDYFH 7
   : |||||
Db 605 NEODYFH 611

RESULT 13
09CU08 PRELIMINARY; PRT; 163 AA.
AC 09CU08;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE 4921540P06RIK PROTEIN (FRAGMENT).
GN 4921540P06RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Knehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staahl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Mordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

```

```

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK015019; BAB2676.1; -.
DR HSSP: P14653; 1B72.
DR MGD: MGI:1913985; 4921540P06RIK.
DR InterPro: IPR001827; Antennapedia.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox_1.
DR PRINTS: PR00025; ANTENNAPEDIA.
DR SMART: SM00389; HOX_1.
DR PROSITE: PS00032; ANTENNAPEDIA; UNKNOWN_1.
DR PROSITE: PS00071; HOMEBOX_2; 1.
FT NON_TER 1
SQ SEQUENCE 163 AA; 18199 MW; E956CF5E76D1E8D1 CRC64;

Query Match          80.0%; Score 32; DB 11; Length 163;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QDYFH 7
   : |||||
Db 11 QDYFH 15

RESULT 14
097SK2 PRELIMINARY; PRT; 180 AA.
AC 097SK2;
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN SP0332.
GN SP0332.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Urdarback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt J.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Mortison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
DR EMBL: AE007345; AKX74507.1; -.
DR TIGR: SP0332; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 180 AA; 20652 MW; E5024C611428B844 CRC64;

Query Match          80.0%; Score 32; DB 16; Length 180;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QDYFH 7
   : |||||
Db 136 QDYFH 140

RESULT 15
P97975

```

ID P97975; PRELIMINARY; PRT; 215 AA.
 AC P97975;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 70-KDA HEAT SHOCK PROTEIN (FRAGMENT).
 GN HSP70.
 OS unidentified soil organism.
 OC unclassified: environmental samples.
 OX NCBI_TaxID=46465;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yap W.H., Li X., Soong T.W., Davies J.E.;
 RT "Genetic diversity of soil microorganisms assessed by analysis of
 RT hsp70 (dnaK) sequences.";
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U49144; AAB48232.1; -.
 DR InterPro: IPR01023; HSP70.
 DR Pfam: PF00012; HSP70.1
 DR PRINTS: PR00301; HEATSHOCK70.
 DR PROSITE: PS00329; HSP70_2; 1.
 DR PROSITE: PS01036; HSP70_3; 1.
 KW Heat shock.
 FT NON_TER 1 1
 FT NON_TER 215 215
 SQ SEQUENCE 215 AA; 23694 MW; 9146A64A3AC9623D CRC64;

Query Match 80.0%; Score 32; DB 14; Length 215;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QDYFH 7
 11111
 Db 196 QDYFH 200

Search completed: June 13, 2002, 09:11:12
 Job time: 1057 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:54:30 ; Search time 57.95 seconds
(without alignments)
2.950 Million cell updates/sec

Title: US-09-710-339-2_COPY_161_167
Perfect score: 40
Sequence: 1 SSQDYFH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	478	1 US-08-720-899-7	Sequence 7, Appl
2	40	100.0	478	1 US-08-459-610-7	Sequence 7, Appl
3	40	100.0	478	2 US-08-343-804-7	Sequence 7, Appl
4	40	100.0	478	2 US-08-339-715A-2	Sequence 2, Appl
5	40	100.0	478	2 US-08-600-908A-10	Sequence 10, Appl
6	40	100.0	478	3 US-08-683-838A-10	Sequence 10, Appl
7	40	100.0	478	3 US-09-182-859-7	Sequence 7, Appl
8	32	80.0	369	1 US-08-700-359-21	Sequence 21, Appl
9	32	80.0	583	2 US-08-616-392C-4	Sequence 4, Appl
10	32	80.0	820	2 US-08-380-182-23	Sequence 23, Appl
11	30	75.0	25	1 US-08-485-455D-41	Sequence 41, Appl
12	30	75.0	25	2 US-08-482-130C-41	Sequence 41, Appl
13	30	75.0	25	2 US-08-484-211C-41	Sequence 41, Appl
14	30	75.0	25	3 US-08-906-769-41	Sequence 41, Appl
15	30	75.0	25	3 US-08-906-616-41	Sequence 41, Appl
16	30	75.0	25	4 US-08-817-795-41	Sequence 41, Appl
17	30	75.0	25	4 US-08-485-443B-41	Sequence 41, Appl
18	30	75.0	25	4 US-08-639-075A-41	Sequence 41, Appl
19	30	75.0	25	4 US-09-012-431-41	Sequence 41, Appl
20	30	75.0	25	4 US-09-012-692-41	Sequence 41, Appl
21	30	75.0	25	4 US-08-906-613-41	Sequence 41, Appl
22	30	75.0	25	5 PCT-US95-14442A-41	Sequence 41, Appl
23	30	75.0	64	1 US-08-485-455D-75	Sequence 75, Appl
24	30	75.0	64	2 US-08-482-130C-75	Sequence 75, Appl
25	30	75.0	64	2 US-08-484-211C-75	Sequence 75, Appl
26	30	75.0	64	3 US-08-906-769-75	Sequence 75, Appl
27	30	75.0	64	3 US-08-906-616-75	Sequence 75, Appl

28	30	75.0	64	4 US-08-817-795-75	Sequence 75, Appl
29	30	75.0	64	4 US-08-485-443B-75	Sequence 75, Appl
30	30	75.0	64	4 US-08-639-075A-75	Sequence 75, Appl
31	30	75.0	64	4 US-09-012-431-75	Sequence 75, Appl
32	30	75.0	64	4 US-09-012-692-75	Sequence 75, Appl
33	30	75.0	64	4 US-08-906-613-75	Sequence 75, Appl
34	30	75.0	64	4 PCT-US95-14442A-75	Sequence 75, Appl
35	30	75.0	223	1 US-08-485-455D-17	Sequence 17, Appl
36	30	75.0	223	2 US-08-482-130C-17	Sequence 17, Appl
37	30	75.0	223	2 US-08-484-211C-17	Sequence 17, Appl
38	30	75.0	223	4 US-08-817-795-17	Sequence 17, Appl
39	30	75.0	223	4 US-08-485-443B-17	Sequence 17, Appl
40	30	75.0	223	5 PCT-US95-14442A-17	Sequence 17, Appl
41	30	75.0	224	3 US-08-906-769-17	Sequence 17, Appl
42	30	75.0	224	3 US-08-906-616-17	Sequence 17, Appl
43	30	75.0	224	4 US-08-639-075A-17	Sequence 17, Appl
44	30	75.0	224	4 US-09-012-431-17	Sequence 17, Appl
45	30	75.0	224	4 US-09-012-692-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
Sequence 7, Application US/08720899
Patent No. 5753460
GENERAL INFORMATION:
APPLICANT: Bisgaard-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
APPLICANT: Svendsen, Allan
APPLICANT: Thellersen, Marianne
APPLICANT: Van der Zee, Pia
TITLE OF INVENTION: AMYLASE VARIANTS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESS: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,899
FILING DATE: 10-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/343,804
FILING DATE: 22-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowrey Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4054,214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-720-899-7
Query Match 100.0%; Score 40; DB 1; Length 478;
Best local Similarity 100.0%; Pred. No 1.9; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

OY 1 SSODYFH 7
|||||
Db 141 SSODYFH 147

RESULT 2
US-08-459-610-7
; Sequence 7, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Franzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 58010430 No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-459-610-7

Query Match 100.0%; Score 40; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
|||||
Db 141 SSODYFH 147

RESULT 3
US-08-343-804-7
; Sequence 7, Application US/08343804
; Patent No. 5850837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Franzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS

NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESS: No. 58308370 No. 5830837disk of No. 5830837th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,804
FILING DATE: 22-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lowney Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4054.214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-343-804-7

Query Match 100.0%; Score 40; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSODYFH 7
|||||
Db 141 SSODYFH 147

RESULT 4
US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michiyo
; APPLICANT: Takata, Hiroki
; APPLICANT: Okada, Shigetaka
; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: POPHAM, HARK, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,715A
; FILING DATE: 14-NOV-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Melser, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335.009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 100.0%; Score 40; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. NO. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
Db 141 SSODYFH 147

RESULT 5
US-08-600-908A-10
Sequence 10, Application US/08600908A
Patent No. 5989169
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 59891690 No. 5989169disk of No. 5989169th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,908A
FILING DATE: 13-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-600-908A-10

Query Match 100.0%; Score 40; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. NO. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
Db 141 SSODYFH 147

RESULT 6
US-08-683-838A-10
Sequence 10, Application US/08683838A
Patent No. 6022724
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 60227240 No. 6022724disk of No. 6022724th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,838A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/600,908
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-683-838A-10

Query Match 100.0%; Score 40; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. NO. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
Db 141 SSODYFH 147

RESULT 7
US-09-182-859-7
Sequence 7, Application US/09182859
Patent No. 6143708
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Borchert, Torben
APPLICANT: Bisgaard-Frantzen, Henrik

TITLE OF INVENTION: Alpha-Amylase Mutants
FILE REFERENCE: 4796.204-US
CURRENT APPLICATION NUMBER: US/09/182,859
CURRENT FILING DATE: 1998-10-29
EARLIER APPLICATION NUMBER: 0515/96
EARLIER FILING DATE: 1996-04-30
EARLIER APPLICATION NUMBER: 0712/96
EARLIER FILING DATE: 1996-06-28
EARLIER APPLICATION NUMBER: 0775/96
EARLIER FILING DATE: 1996-07-11
EARLIER APPLICATION NUMBER: 1263/96
EARLIER FILING DATE: 1996-11-08
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 7
LENGTH: 478
TYPE: PRT
ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match 100.0%; Score 40; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSODYFH 7
DB 141 SSODYFH 147

RESULT 8
US-08-700-359-21
Sequence 21, Application US/08700359
Patent No. 5766925
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, MASAKAZU
APPLICANT: USUDA, YOSHIHIRO
APPLICANT: SUZUKI, TOMOKO
APPLICANT: TANAKA, AKIKO
APPLICANT: MATSUI, HIROSHI
TITLE OF INVENTION: METHOD OF PRODUCING L-LYSINE
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: OHLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS TEXT EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,359
FILING DATE: 08-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-35019
FILING DATE: 04-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OHLON
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-819-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
TOPOLOGY: linear

US-08-700-359-21

Query Match 80.0%; Score 32; DB 1; Length 369;
Best Local Similarity 71.4%; Pred. No. 54;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSODYFH 7
DB 133 SSMDYFH 139

RESULT 9
US-08-616-392C-4
Sequence 4, Application US/08616392C
Patent No. 5998165
GENERAL INFORMATION:
APPLICANT: Goold, Richard D.
APPLICANT: Akerblom, Ingrid E.
APPLICANT: Sellhammer, Jeffrey
APPLICANT: Coleman, Roger
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES PANCLIA
TITLE OF INVENTION: AND PANCLIA ASSOCIATED WITH PANCREATIC CANCER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,392C
FILING DATE: 15-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/581,240
FILING DATE: 29-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0052-105
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: hnt
CLONE: 436071
US-08-616-392C-4

Query Match 80.0%; Score 32; DB 2; Length 583;
Best Local Similarity 71.4%; Pred. No. 87;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSODYFH 7
DB 433 STDDYFH 439

RESULT 10
US-08-380-182-23
Sequence 23, Application US/08380182

```

; Patent No. 5858749
; GENERAL INFORMATION:
; APPLICANT: Matthews, Benjamin F.
; APPLICANT: Weisemann, Jane M.
; TITLE OF INVENTION: A Bifunctional Protein From Carrots
; TITLE OF INVENTION: (Daucus carota) with Aspartokinase and Homoserine
; TITLE OF INVENTION: Dehydrogenase Activities"
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Janelle S. Graeter
; STREET: Room 411, Bldg. 005, BARC-W
; CITY: Beltsville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20705
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/380,182
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Graeter, Janelle S.
; REGISTRATION NUMBER: 35,024
; REFERENCE/DOCKET NUMBER: 0226.94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-504-6629
; TELEFAX: 301-504-5060
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 820 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
;
US-08-380-182-23

```

```

Query Match      80.0%; Score 32; DB 2; Length 820;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 SSODYFH 7
Db 584 SSMDYFH 590

```

```

RESULT 11
US-08-485-455D-41
; Sequence 41, Application US/08485455D
; Patent No. 5712143
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary L.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, AND USES THEREOF
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,455D
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C1-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-485-455D-41

```

```

Query Match      75.0%; Score 30; DB 1; Length 25;
Best Local Similarity 66.7%; Pred. No. 8;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 2 SQDYFH 7
Db 20 AEDYFH 25

```

```

RESULT 12
US-08-482-130C-41
; Sequence 41, Application US/08482130C
; Patent No. 5962257
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary L.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, AND USES THEREOF
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,130C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C1-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-130C-41

Query Match
Best Local Similarity 75.0%; Score 30; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 8;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 SODYFH 7
:|||||
DB 20 AEDYFH 25

RESULT 13
US-08-484-211C-41
Sequence 41, Application US/08484211C
Patent No. 5972645
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary L.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,211C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-211C-41

Query Match
Best Local Similarity 75.0%; Score 30; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 8;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 SODYFH 7
:|||||
DB 20 AEDYFH 25

RESULT 14

US-08-906-769-41
Sequence 41, Application US/08906769
Patent No. 6077687
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary L.
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,769
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-906-769-41

Query Match
Best Local Similarity 75.0%; Score 30; DB 3; Length 25;
Best Local Similarity 66.7%; Pred. No. 8;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 SODYFH 7
:|||||
DB 20 AEDYFH 25

RESULT 15
US-08-906-616-41
Sequence 41, Application US/08906616
Patent No. 6121035
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary L.
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
NUMBER OF SEQUENCES: 190

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; City: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,616
; FILING DATE: 05-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-906-616-41

Query Match 75.0%; Score 30; DB 3; Length 25;
Best Local Similarity 66.7%; Pred No. 8;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 SQDYFH 7
: : : : :
Db 20 AEDYFH 25

Search completed: June 13, 2002, 08:54:31
Job time: 316 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:17 ; Search time 125.81 Seconds

(without alignments)
11.000 Million cell updates/sec

Title: US-09-710-339-2_COPY_468_475

Perfect score: 45
Sequence: 1 DGNVPPM 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	45	100.0	499	3	Q96TH4	Q96TH4 aspergillus
2	37	82.2	6260	2	Q54299	Q54299 streptomyc
3	36	80.0	335	4	Q9BWN2	Q9BWN2 homo sapien
4	36	80.0	360	4	Q9BQ81	Q9BQ81 homo sapien
5	36	80.0	457	16	Q987U9	Q987U9 rhizobium 1
6	36	80.0	619	5	Q9VKJ8	Q9VKJ8 drosophila
7	36	80.0	1385	12	Q65565	Q65565 bovine heip
8	35	77.8	98	5	Q9U6B4	Q9U6B4 echinometra
9	35	77.8	612	10	Q81829	Q81829 arabidopsis
10	35	77.8	650	10	Q94209	Q94209 oryza sativ
11	35	77.8	652	12	Q10436	Q10436 mouse adeno
12	35	77.8	722	2	P96091	P96091 treponema d
13	35	77.8	1392	12	Q9E1X9	Q9E1X9 cercopithec
14	35	77.8	1989	11	Q9ERT7	Q9ERT7 mus musculu
15	35	77.8	8563	2	Q54297	Q54297 streptomyc
16	35	77.8	10223	2	Q54296	Q54296 streptomyc

17	34	75.6	297	16	Q928V2	Q928V2 chlamydia p
18	34	75.6	335	16	Q98KV6	Q98KV6 rhizobium 1
19	34	75.6	465	2	Q69840	Q69840 streptomyc
20	34	75.6	472	11	Q9DAP3	Q9DAP3 mus musculu
21	34	75.6	585	5	Q9V36	Q9V36 drosophila
22	34	75.6	593	10	Q93VE2	Q93VE2 oryza sativ
23	34	75.6	640	3	Q13296	Q13296 aspergillus
24	34	75.6	732	5	Q9GR74	Q9GR74 leishmania
25	34	75.6	989	16	Q9CNT9	Q9CNT9 pasteurella
26	34	75.6	1390	10	Q49145	Q49145 arabidopsis
27	33.5	74.4	90	16	Q31786	Q31786 bacillus su
28	33	73.3	111	13	Q9DDT3	Q9DDT3 atalinga mi
29	33	73.3	111	13	Q9DDT1	Q9DDT1 ara araraun
30	33	73.3	125	17	Q59155	Q59155 pyrococcus
31	33	73.3	238	2	Q91LA9	Q91LA9 streptomyc
32	33	73.3	261	5	Q9XUR6	Q9XUR6 caenorhabdi
33	33	73.3	271	17	Q9HQQ3	Q9HQQ3 thermoplasm
34	33	73.3	288	17	Q96XP7	Q96XP7 sulfolobus
35	33	73.3	362	11	Q9CX00	Q9CX00 mus musculu
36	33	73.3	372	3	Q91BH2	Q91BH2 tricholoma
37	33	73.3	373	10	Q9FUC4	Q9FUC4 nicotiana t
38	33	73.3	403	16	Q9KAV1	Q9KAV1 bacillus ha
39	33	73.3	451	2	Q05074	Q05074 streptomyc
40	33	73.3	507	17	Q52024	Q52024 halobacteri
41	33	73.3	525	5	Q97229	Q97229 plasmodium
42	33	73.3	529	5	Q962K9	Q962K9 plasmodium
43	33	73.3	607	4	Q9NVM2	Q9NVM2 homo sapien
44	33	73.3	637	16	Q985R6	Q985R6 rhizobium 1
45	33	73.3	644	16	Q985R7	Q985R7 rhizobium 1

ALIGNMENTS

RESULT 1
ID Q96TH4 PRELIMINARY: PRT: 499 AA.
AC Q96TH4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TRKA-AMYLASE A (EC 3.2.1.1) (ALPHA-AMYLASE).
GN AMYA OR AMY1.
OS Aspergillus oryzae, and
OS Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062, 5059;
RP SEQUENCE FROM N.A.
RC SPECIES=A.oryzae; STRAIN=RTB40;
RX MEDLINE=20289310; PubMed=10830498;
RA Gomi K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N.,
RA Iimura Y.;
RT "Molecular cloning and characterization of a transcriptional activator
RT gene, amyR, involved in the amylolytic gene expression in Aspergillus
RT oryzae".
RT Biosci. Biotechnol. Biochem. 64:816-827(2000).
RL [2]
RN SEQUENCE FROM N.A.
RP SPECIES=A.flavus; STRAIN=86-10D;
RA Pakhoury A.M., Woloshuk C.P.;
RT "AmyR, the alpha-amylase gene of Aspergillus flavus: Involvement in
RT aflatoxin biosynthesis in maize kernels".
RL Phytopathology 89:908-914(1999).
DR EMBL: AB021876; BAA95703.1; -;
DR EMBL: AF139925; AAF14264.1; -;
SQ SEQUENCE 499 AA; 54804 MW; EEF42ADA7ID20DA9 CRC64;

Query Match 100.0%; Score 45; DB 3; Length 499;
Best local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
 DB 469 DGNVPVPM 476

RESULT 2
 ID 054299 PRELIMINARY; PRT; 6260 AA.
 AC 054299;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE POLYKETIDE SYNTHASE.
 GN RAPC.
 OS Streptomyces hygroscopicus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1912;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NRRL 5491;
 RC MEDLINE=95372374; PubMed=7644502;
 RA Schwecke T., Aparicio J.F., Molnar I., Koenig A., Khaw L.E.,
 RA Haydock S.F., O'Leary M., Caffrey P., Cortes J., Lester J.B.,
 RA Boehm G.A., Staunton J., Leadlay P.F.;
 RT "The biosynthetic gene cluster for the polyketide immunosuppressant
 rapamycin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:7839-7843(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NRRL 5491;
 RC MEDLINE=96186895; PubMed=8635730;
 RA Molnar I., Aparicio J.F., Haydock S.F., Khaw L.E., Schwecke T.,
 RA Koenig A., Staunton J., Leadlay P.F.;
 RT "Organisation of the biosynthetic gene cluster for rapamycin in
 Streptomyces hygroscopicus: analysis of genes flanking the polyketide
 synthase.";
 RL gene 169:1-7(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NRRL 5491;
 RC MEDLINE=96186896; PubMed=8635756;
 RA Aparicio J.F., Molnar I., Schwecke T., Koenig A., Haydock S.F.,
 RA Khaw L.E., Staunton J., Leadlay P.F., Staunton J.,
 RT "Organisation of the biosynthetic gene cluster for rapamycin in
 Streptomyces hygroscopicus: analysis of the enzymatic domains in the
 modular polyketide synthase.";
 RL gene 169:9-16(1996).
 DR EMBL: X86780; CAA60462.1; -.
 DR HSSP: P25715; 1MLA.
 DR InterPro: IPR001227; Acyltransferase domain.
 DR InterPro: IPR002085; Adh_zn_family.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR000794; ketoacyl-synth.
 DR InterPro: IPR003880; Phosphopant_c_4.
 DR InterPro: IPR002364; QOR_zeta_crystal.
 DR Pfam: PF00698; Acyl_transf. 4.
 DR Pfam: PF00107; adh_zinc. 1.
 DR Pfam: PF00109; ketoacyl-synth. 4.
 DR Pfam: PF02801; ketoacyl-synth. C. 4.
 DR Pfam: PF00550; pp-binding. 4.
 DR PROSITE: PS50075; ACP_DOMAIN; 4.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 4.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; 4.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
 DR PROSITE: PS01162; QOR_ZETA_CRYSTAL; 1.
 DR Phosphopantetheine; Transferase.
 SO SEQUENCE 6260 AA; 659509 MW; D7D6C002F91F2080 CRC64;

Query Match 82.2%; Score 37; DB 2; Length 6260;
 Best Local Similarity 85.7%; Pred. No. 7e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 DGNVPVPM 7
 DB 4372 DGNVPVPM 4378

RESULT 3
 ID 09BWN2 PRELIMINARY; PRT; 335 AA.
 AC 09BWN2;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE SIMILAR TO KIAA0174 GENE PRODUCT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EYE, RETINOBLASTOMA;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC000116; AAH00116.1; -.
 SO SEQUENCE 335 AA; 36622 MW; 7D0714269380E835 CRC64;

Query Match 80.0%; Score 36; DB 4; Length 335;
 Best Local Similarity 75.0%; Pred. No. 41;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
 DB 225 DGTVPVPM 232

RESULT 4
 ID 09B081 PRELIMINARY; PRT; 360 AA.
 AC 09B081;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE SIMILAR TO KIAA0174 GENE PRODUCT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LUNG CARCINOMA;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LUNG CARCINOMA;
 RC Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC000435; AAH04359.1; -.
 DR EMBL: BC000430; AAH00430.1; -.
 SO SEQUENCE 360 AA; 39927 MW; 5D0DD8A113DAA2BC CRC64;

Query Match 80.0%; Score 36; DB 4; Length 360;
 Best Local Similarity 75.0%; Pred. No. 45;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
 DB 225 DGTVPVPM 232


```

RESULT      5
Q98709      PRELIMINARY;      PRT;      457 AA.
AC Q98709;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE GMMMA-GLUTAMYL-STEINE SYNTHETASE.
GN MLI6902.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC PHYLIobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFE30309;
RX Kameko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Kameko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003010; BAB53101.1; -.
KW Complete proteome.
SQ SEQUENCE 457 AA; 51437 MW; 9DFE7989DD10CF38 CRC64;

```

```

Query Match      80.0%; Score 36; DB 16; Length 457;
Best Local Similarity 85.7%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 DGNPVP 7
DB 45 DGNAPVP 51

```

```

RESULT      6
Q9VKJ8      PRELIMINARY;      PRT;      619 AA.
ID Q9VKJ8;
AC Q9VKJ8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG6201 PROTEIN.
GN CG6201.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRKRELEV;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abil J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu U., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Bottlier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

```

```

RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Modarity C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skipski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003631; AAF53070.1; -.
DR FLYBase: FBgn0032343; CG6201.
DR InterPro: IPR000514; Glyco_hydro.39.
DR Pfam: PF01229; Glyco_hydro.39; 1.
DR PRINTS: PR00745; GHYDRASE39.
SQ SEQUENCE 619 AA; 70341 MW; A7A79C7BF93A14DA CRC64;

```

```

Query Match      80.0%; Score 36; DB 5; Length 619;
Best Local Similarity 75.0%; Pred. No. 83;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 DGNVPPM 8
DB 207 DGKVLPM 214

```

```

RESULT      7
Q65565      PRELIMINARY;      PRT;      1385 AA.
ID Q65565;
AC Q65565;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MAJOR CAPSID PROTEIN.
GN UL19.
OS Bovine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COOPER, AND JURA;
RX MEDLINE=95313343; PubMed=7793062;
RA Vilek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
RA Lechworth G.J., Schwzyer M.;
RT "Nucleotide sequence analysis of a 30-kb region of the bovine
RT herpesvirus 1 genome which exhibits a colinear gene arrangement with
RT the UL21 to UL4 genes of herpes simplex virus."
RL Virology 210:100-108(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=JURA;
RA Schwzyer M., Paces V., Letchworth G.J., Mista V., Buhk H.J.,
RA Lowery D.E., Simard C., Bello L.J., Thiry E., Vilek C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=JURA;
RA Schwzyer M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; Z48053; CAA8914.1; -.
 DR EMBL; AJ004801; CAA06115.1; -.
 DR InterPro; IPR000912; Herpes_MCP.
 DR Pfam; PF03122; Herpes_MCP.1.
 DR PRINTS; PR00235; HSVCAPSIDMCP.
 SO SEQUENCE 1385 AA; 149735 MW; 1B2117B7877060A3 CRC64;

Query Match 80.0%; Score 36; DB 12; Length 1385;
 Best Local Similarity 62.5%; Pred. No. 2e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DGNVPVP 8
 Db 580 GNTVPPL 587

RESULT 8
 O906A PRELIMINARY; PRT; 98 AA.
 AC O906A;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE BINDIN (FRAGMENT).
 OS Echinomembra mathaei.
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoida; Euechinoida; Echinacea; Echinoida; Echinometridae;
 OC Echinometra.
 OK NCBI_TaxID=31178;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M6B.1;
 RX MEDLINE=20006286; PubMed=10535974;
 RA Palumbi S.R.;
 RT "All males are not created equal: fertility difference depend on
 gamete recognition polymorphisms in sea urchins";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:12632-12637(1999).
 DR EMBL; AF186336; AAF07150.1; -.
 DR InterPro; IPR000775; Bindin.
 DR Pfam; PF02084; Bindin; 1.
 FT NON_TER 1 98
 FT NON_TER 98 1
 SO SEQUENCE 98 AA; 9660 MW; 579B07577FB70EFC CRC64;

Query Match 77.8%; Score 35; DB 5; Length 98;
 Best Local Similarity 85.7%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 DGNVPVP 7
 Db 14 DGNVPVP 20

RESULT 9
 O81829 PRELIMINARY; PRT; 612 AA.
 AC O81829;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE G3 LIKE PROTEIN.
 GN M4122.70 OR ATAG27260.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OK NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Reichert B.J., Barel E., Hohnsels J., Mewes H.W., Mayer K.,
 RA Scheller C., Bevan M.;

RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Reichert B.J., Barel E., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AL030978; CAA19720.1; -.
 DR EMBL; AL161566; CAB79581.1; -.
 SO SEQUENCE 612 AA; 69283 MW; 8EDF91FE7EE2725A CRC64;

Query Match 77.8%; Score 35; DB 10; Length 612;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 DGNVPVP 7
 Db 506 GNTVPPL 512

RESULT 10
 O94209 PRELIMINARY; PRT; 650 AA.
 AC O94209;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUTATIVE POTASSIUM TRANSPORT PROTEIN.
 GN P0506E04.23.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OK NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0506E04";
 RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AP003272; BAB67845.1; -.
 SO SEQUENCE 650 AA; 70963 MW; 9BDBC4F6870B4B50 CRC64;

Query Match 77.8%; Score 35; DB 10; Length 650;
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GNPVPVP 8
 Db 160 GNTVPPL 166

RESULT 11
 ID 010436 PRELIMINARY; PRT; 652 AA.
 AC 010436;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PTP (FRAGMENT).
 GN E2B PTP.
 OS Mouse adenovirus type 1 (MAV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OK NCBI_TaxID=10530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Meissner J.D., Hirsch G.N., Larue E.A., Fulcher R.A., Spindler K.R.;
 RT "Completion of the DNA sequence of mouse adenovirus type 1: sequence

RT of E2B, L1, and L2 (18-51 map units)";
 RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U93843; AAB53751.1; -;
 DR InterPro: IPR003391; Adeno_terminal.
 DR Pfam: PF02459; Adeno_terminal; 1.
 FT NON_TER 1
 FT VARIANT 644 644 E -> G.
 SQ SEQUENCE 652 AA; 74664 MW; 1400CF2D2E26DD0B CRC64;

Query Match 77.8%; Score 35; DB 12; Length 652;
 Best Local Similarity 75.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
 ||| |||
 Db 202 DGNVPVP 209

RESULT 12
 ID P96091 PRELIMINARY; PRT; 722 AA.
 AC P96091;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SERINE PROTEASE (DENTILISIN) FROM TREPONEMA DENTICOLA PRECURSOR.
 GN PRTP.
 OS Treponema denticola.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_TaxID=158;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX MEDLINE=97101036; PubMed=8945563;
 RA Ishihara K., Miura T., Kuramitsu H.K., Oda K.;
 RT "Characterization of the Treponema denticola prtp gene encoding a
 RT prolyl-phenylalanine-specific protease (dentilisins).";
 RL Infect. Immun. 64:5178-5186(1996).
 DR EMBL: D83264; BAA11874.1; -;
 DR HSSP: O45670; IDBI.
 DR MEROPS: S08.024; -;
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 2.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KM Signal: Protease.
 FT SIGNAL 1
 FT SIGNAL 25 POTENTIAL.
 SQ SEQUENCE 722 AA; 77474 MW; A604DA078646910D CRC64;

Query Match 77.8%; Score 35; DB 2; Length 722;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVP 7
 ||| |||
 Db 511 DGNVPVP 517

RESULT 13
 ID Q9EIX9 PRELIMINARY; PRT; 1392 AA.
 AC Q9EIX9;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MAJOR CAPSID PROTEIN.
 OS Cercarial dermatitis virus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.

OX NCBI_TaxID=35245;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;
 RT "Complete Sequence of the Simian Varicella Virus Genome.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF275348; AAC27215.1; -;
 DR InterPro: IPR000912; Herpes_MCP.
 DR PRINTS: PR00235; HSVCAPSIDMCP.
 SQ SEQUENCE 1392 AA; 155857 MW; C94599418AF62E19 CRC64;

Query Match 77.8%; Score 35; DB 12; Length 1392;
 Best Local Similarity 50.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
 .||| |||
 Db 591 GNIPIPL 598

RESULT 14
 ID Q9ERT7 PRELIMINARY; PRT; 1989 AA.
 AC Q9ERT7;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE SYNE-1B (FRAGMENT).
 GN SYNE1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX MEDLINE=20493380; PubMed=10878022;
 RA Apel E.D., Lewis R.M., Grady R.M., Sanes J.R.;
 RT "Syne-1, a dystrophin- and klarsicht-related protein associated with
 RT synaptic nuclei at the neuromuscular junction.";
 RL J. Biol. Chem. 275:31986-31995(2000).
 DR EMBL: AF281870; AAC24393.1; -;
 DR MGD: MGI:1927152; Syne1.
 DR InterPro: IPR002017; Spectrin.
 DR Pfam: PF00435; Spectrin; 14.
 DR SMART: SM00150; Spec; 14.
 FT NON_TER 1
 SQ SEQUENCE 1989 AA; 229066 MW; 381E6AF211CC138D CRC64;

Query Match 77.8%; Score 35; DB 11; Length 1989;
 Best Local Similarity 71.4%; Pred. No. 4.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 GNVVPVP 8
 ||| |||
 Db 786 GNIPIPL 792

RESULT 15
 ID Q54297 PRELIMINARY; PRT; 8563 AA.
 AC Q54297;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE POLYKETIDE SYNTHASF.
 GN RAPA.
 OS Streptomyces hygroscopicus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1912;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN-NRRL 5491;
RX MEDLINE=95372374; PubMed=7644502;
RA Schwecke T., Aparicio J.F., Molnar I., Koenig A., Khaw L.E.,
RA Haydock S.F., Oillynk M., Caffrey P., Cortes J., Lester J.B.,
RA Boehm G.A., Staunton J., Leadlay P.F.;
RT "The biosynthetic gene cluster for the polypeptide immunosuppressant
RT rapamycin.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:7839-7843(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NRRL 5491;
RX MEDLINE=96186895; PubMed=8635730;
RA Molnar I., Aparicio J.F., Haydock S.F., Khaw L.E., Schwecke T.,
RA Koenig A., Staunton J., Leadlay P.F.;
RT "Organisation of the biosynthetic gene cluster for rapamycin in
RT Streptomyces hygroscopicus: analysis of genes flanking the polypeptide
RT synthase.";
RL Gene 169:1-7(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-NRRL 5491;
RX MEDLINE=96186895; PubMed=8635756;
RA Aparicio J.F., Molnar I., Schwecke T., Koenig A., Haydock S.F.,
RA Khaw L.E., Staunton J., Leadlay P.F., Staunton J., Leadlay P.F.;
RT "Organization of the biosynthetic gene cluster for rapamycin in
RT Streptomyces hygroscopicus: analysis of the enzymatic domains in the
RT modular polypeptide synthase.";
RL Gene 169:9-16(1996).
DR EMBL; X86780; CAA60460.1; -.
DR HSSP; P08659; ILCI.
DR InterPro: IPR001227; Acyltransf_domain.
DR InterPro: IPR002085; Adh_zn_family.
DR InterPro: IPR000873; AMP-bind.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR000794; Ketoacyl-synt.
DR InterPro: IPR003880; Phosphopant_attach.
DR InterPro: IPR002364; OOR_zeta_crystal.
DR Pfam; PF00698; Acyl-transf; 4.
DR Pfam; PF00107; adh_zinc; 3.
DR Pfam; PF00501; AMP-binding; 1.
DR Pfam; PF00109; ketoacyl-synt; 4.
DR Pfam; PF02801; ketoacyl-synt_C; 4.
DR Pfam; PF00550; pp-binding; 5.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS50075; ACP_DOMAIN; 5.
DR PROSITE; PS00455; AMP_BINDING; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 4.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_3.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 5.
DR PROSITE; PS01162; OOR_ZETA_CRISTAL; 2.
DR Phosphatetheine; transferase.
SQ SEQUENCE 8563 AA; 899948 MW; FEA462CD33838A37 CRC64;

```

```

Query Match 77.88; Score 35; DB 2; Length 8563;
Best Local Similarity 85.78; Pred. NO. 2.5e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 DGNVPP 7
DB 4389 DGDVPP 4395

```

Search completed: June 13, 2002, 09:11:19
Job time: 1064 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:54:33 ; Search time 57.95 Seconds
(without alignments)
3.372 Million cell updates/sec

Title: US-09-710-339-2_COPY_468_475

Perfect score: 45

Sequence: 1 DGNVPPM 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/prodata/2/1aa/CTUS.COMB.pep: *
6: /cgn2_6/prodata/2/1aa/Backfilest.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	478	1	US-08-720-899-7
2	45	100.0	478	2	US-08-459-610-7
3	45	100.0	478	2	US-08-343-804-7
4	45	100.0	478	2	US-08-339-715A-2
5	45	100.0	478	2	US-08-600-908A-10
6	45	100.0	478	3	US-08-683-838A-10
7	45	100.0	478	4	US-09-182-859-7
8	32	71.1	14	5	US-08-408-604A-15
9	32	71.1	14	5	PCT-US93-09626-15
10	32	71.1	230	3	US-09-167-434-2
11	32	71.1	230	3	US-08-853-755-2
12	31	68.9	261	4	US-09-045-973-3
13	30	66.7	40	2	US-08-448-481-4
14	30	66.7	45	1	US-08-217-687-41
15	30	66.7	45	2	US-08-777-192-41
16	30	66.7	45	4	US-08-971-982-41
17	30	66.7	53	2	US-08-640-847C-40
18	30	66.7	94	1	US-08-591-498-4
19	30	66.7	112	1	US-08-326-117B-17
20	30	66.7	112	3	US-08-982-129-17
21	30	66.7	338	4	US-09-413-231-6
22	30	66.7	338	4	US-09-413-231-7
23	30	66.7	449	5	US-08-657-393-2
24	30	66.7	449	5	PCT-US94-02539-2
25	30	66.7	459	2	US-08-673-312-2
26	30	66.7	471	2	US-08-657-392-31
27	30	66.7	471	5	PCT-US94-02539-31

28	30	66.7	589	2	US-08-756-317-5	Sequence 5, Appl
29	30	66.7	711	1	US-08-235-838-7	Sequence 7, Appl
30	30	66.7	711	2	US-08-465-473B-7	Sequence 4, Appl
31	30	66.7	893	4	US-09-514-302-4	Sequence 12, Appl
32	30	66.7	907	4	US-08-989-299-12	Sequence 3, Appl
33	30	66.7	1253	4	US-07-920-281C-3	Sequence 2, Appl
34	30	66.7	1528	1	US-08-466-277-3	Sequence 2, Appl
35	30	66.7	1528	3	US-08-326-117B-2	Sequence 2, Appl
36	30	66.7	1528	3	US-08-982-129-2	Sequence 2, Appl
37	30	66.7	1938	4	US-09-514-302-2	Sequence 2, Appl
38	29	64.4	33	4	US-08-810-712-30	Sequence 30, Appl
39	29	64.4	337	1	US-08-440-856A-3	Sequence 12, Appl
40	29	64.4	415	1	US-08-451-715A-12	Patent No. 5194375
41	29	64.4	459	6	5194375-6	Sequence 2, Appl
42	29	64.4	524	1	US-08-461-837-2	Sequence 12, Appl
43	29	64.4	524	2	US-08-480-736-12	Sequence 2, Appl
44	29	64.4	524	3	US-08-973-223-2	Sequence 2, Appl
45	29	64.4	524	4	US-09-347-060-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-720-899-7
Sequence 7, Application US/08720899
Patent No. 5753460
GENERAL INFORMATION:
APPLICANT: Bisgaard-Frantzen, Henrik
APPLICANT: Borchart, Torben Vedel
APPLICANT: Svendsen, Allan
APPLICANT: Thøgersen, Marianne
APPLICANT: Van der Zee, Pia
TITLE OF INVENTION: ARTILASE VARIANTS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5753460a No. 5753460a of No. 5753460a America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08720, 899
FILING DATE: 10-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/343, 804
FILING DATE: 22-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowrey Dr. Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4054.214-US
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-720-899-7

Query Match 100.0%; Score 45; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.4; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

OY 1 DGNVPPM 8
 |||||
 Db 448 DGNVPPM 455

RESULT 2
 US-08-459-610-7
 ; Sequence 7, Application US/08459610
 ; Patent No. 5801043
 ; GENERAL INFORMATION:
 ; APPLICANT: Bisgaard-Frantzen, Henrik
 ; APPLICANT: Borchert, Torben Vedel
 ; APPLICANT: Svendsen, Allan
 ; APPLICANT: Thellersen, Marianne
 ; TITLE OF INVENTION: AMYLASE VARIANTS
 ; NUMBER OF SEQUENCES: 38
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 5801043 of No. 5801043th America, Inc.
 ; STREET: 405 Lexington Avenue, 64th Floor
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10174-6401
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/459,610
 ; FILING DATE: 02-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/343,804
 ; FILING DATE: 22-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney Dr., Karen A.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 4054, 214-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 478 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-459-610-7

Query Match 100.0%; Score 45; DB 1; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPPM 8
 |||||
 Db 448 DGNVPPM 455

RESULT 3
 US-08-343-804-7
 ; Sequence 7, Application US/08343804
 ; Patent No. 5830837
 ; GENERAL INFORMATION:
 ; APPLICANT: Bisgaard-Frantzen, Henrik
 ; APPLICANT: Borchert, Torben Vedel
 ; APPLICANT: Svendsen, Allan
 ; APPLICANT: Thellersen, Marianne
 ; APPLICANT: Van der Zee, Pia
 ; TITLE OF INVENTION: AMYLASE VARIANTS

NUMBER OF SEQUENCES: 38
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 5830837 of No. 5830837th America, Inc.
 ; STREET: 405 Lexington Avenue, 64th Floor
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10174-6401
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/343,804
 ; FILING DATE: 22-NOV-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney Dr., Karen A.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 4054, 214-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 478 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-343-804-7

Query Match 100.0%; Score 45; DB 2; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPPM 8
 |||||
 Db 448 DGNVPPM 455

RESULT 4
 US-08-339-715A-2
 ; Sequence 2, Application US/08339715A
 ; Patent No. 5965442
 ; GENERAL INFORMATION:
 ; APPLICANT: Kaneko, Hiroki
 ; APPLICANT: Takada, Toshikazu
 ; APPLICANT: Shimada, Jiro
 ; APPLICANT: Kuriki, Takashi
 ; APPLICANT: Yanase, Michiyo
 ; APPLICANT: Takata, Hiroki
 ; APPLICANT: Okada, Shigetaka
 ; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
 ; TITLE OF INVENTION: NEOPULULANASE
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: POPHAM, HAIK, SCHNOBRICH & KAUFMAN, LTD.
 ; STREET: Metropolitan Square Building, Suite 800, 1450
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/339,715A
 ; FILING DATE: 14-NOV-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Malsier, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335.009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 100.0%; Score 45; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DGNVPVPM 8
DB 447 DGNVPVPM 454

RESULT 5
US-08-600-908A-10
Sequence 10, Application US/08600908A
Patent No. 5989169
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59891690 No. 5989169disk of No. 5989169th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,908A
FILING DATE: 13-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-600-908A-10

Query Match 100.0%; Score 45; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DGNVPVPM 8
DB 448 DGNVPVPM 455

RESULT 6
US-08-683-838A-10
Sequence 10, Application US/08683838A
Patent No. 6022724
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60227240 No. 6022724disk of No. 6022724th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,838A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/600,908
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-683-838A-10

Query Match 100.0%; Score 45; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DGNVPVPM 8
DB 448 DGNVPVPM 455

RESULT 7
US-09-182-859-7
Sequence 7, Application US/09182859
Patent No. 6143708
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Borchert, Torben
APPLICANT: Bisgard-Frantzen, Henrik

;; TITLE OF INVENTION: Alpha-Amylase Mutants
;; FILE REFERENCE: 4796.204-US
;; CURRENT APPLICATION NUMBER: US/09/182,859
;; CURRENT FILING DATE: 1998-10-29
;; EARLIER APPLICATION NUMBER: 0515/96
;; EARLIER FILING DATE: 1996-04-30
;; EARLIER APPLICATION NUMBER: 0712/96
;; EARLIER FILING DATE: 1996-06-28
;; EARLIER APPLICATION NUMBER: 0775/96
;; EARLIER FILING DATE: 1996-07-11
;; EARLIER APPLICATION NUMBER: 1263/96
;; EARLIER FILING DATE: 1996-11-08
;; NUMBER OF SEQ ID NOS: 37
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 7
;; LENGTH: 478
;; TYPE: PRT
;; ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match 100.0%; Score 45; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 1,4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
DB 448 DGNVPVPM 455

RESULT 8
US-08-408-604A-15
; Sequence 15, Application US/08408604A
; Patent No. 5801149
; GENERAL INFORMATION:
; APPLICANT: SHOELSON, Steven
; TITLE OF INVENTION: INHIBITION OF SIGNAL TRANSDUCTION MOLECULES
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/408,604A
; FILING DATE: 21-MAR-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/134,558
; FILING DATE: 08-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/959,949
; FILING DATE: 09-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/722,359
; FILING DATE: 19-JUNE-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: JDP-014CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14

;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-408-604A-15

Query Match 71.1%; Score 32; DB 1; Length 14;
Best Local Similarity 62.5%; Pred. No. 6;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
DB 3 DGNVPVPM 10

RESULT 9
PCT-US93-09626-15
; Sequence 15, Application PC/TUS9309626
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: INHIBITION OF SIGNAL TRANSDUCTION MOLECULES
; NUMBER OF SEQUENCES: 48
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/09626
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/959,949
; FILING DATE: October 9, 1992
; APPLICATION NUMBER: 07/722,359
; FILING DATE: June 19, 1991
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US93-09626-15

Query Match 71.1%; Score 32; DB 5; Length 14;
Best Local Similarity 62.5%; Pred. No. 6;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
DB 3 DGNVPVPM 10

RESULT 10
US-09-167-434-2
; Sequence 2, Application US/09167434
; Patent No. 6008033
; GENERAL INFORMATION:
; APPLICANT: Abdel-Meguid, Sherin
; APPLICANT: Qiu, Xiyang
; APPLICANT: Culp, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Janson, Cheryl A.
; APPLICANT: Hoog, Susan S.
; APPLICANT: Smith, Ward W.
; TITLE OF INVENTION: No. 6008033tel Proteases, Compositions Capable of
; TITLE OF INVENTION: Binding to Said Site, and Methods of Use Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road - P.O. Box 1539
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA


```

;
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,434
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/853,755
; FILING DATE:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/022,470
; FILING DATE: 26-JUL-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,416
; FILING DATE: 21-AUG-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,901
; FILING DATE: 14-NOV-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/035,973
; FILING DATE: 21-JAN-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,191
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Diner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: P50472-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5022
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-167-434-2

Query Match 71.1%; Score 32; DB 3; Length 230;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 GNPVPM 8
DB 36 GNIPPL 42

RESULT 11
; Sequence 2, Application US/08853755
; Patent No. 6083711
; GENERAL INFORMATION:
; APPLICANT: Abdel-Meguid, Sherin
; APPLICANT: Qiu, Xiyang
; APPLICANT: Culp, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Janson, Cheryl A.
; APPLICANT: Hoog, Susan S.
; APPLICANT: Smith, Ward W.
; TITLE OF INVENTION: No. 6083711el Proteases, Compositions Capable of
; TITLE OF INVENTION: Binding to Said Site, and Methods of Use Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road - P.O. Box 1539
; CITY: King of Prussia
```

```

;
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,755
; FILING DATE:
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/018,616
; FILING DATE: 15-MAY-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/022,470
; FILING DATE: 26-JUL-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,416
; FILING DATE: 21-AUG-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,901
; FILING DATE: 14-NOV-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/035,973
; FILING DATE: 21-JAN-1997
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,191
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Diner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: P50472-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5022
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-853-755-2

Query Match 71.1%; Score 32; DB 3; Length 230;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 GNPVPM 8
DB 36 GNIPPL 42

RESULT 12
; Sequence 3, Application US/09045973
; Patent No. 6165767
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Maiah
; TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
```

COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,973
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0491 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 261 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT18
CLONE: 2534680
US-09-045-973-3

Query Match 68.9%; Score 31; DB 4; Length 261;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVFP 7
111111
DB 70 DGAVIDP 76

RESULT 13
US-08-448-481-4
Sequence 4, Application US/08448481
Patent No. 5814270
GENERAL INFORMATION:
APPLICANT: Coulos-Thyenot, Pierre M.L.
APPLICANT: Jouenne, Thierry G.
APPLICANT: Maes, Olivier C.A.
APPLICANT: Deloire, Alain J.
APPLICANT: Boulay, Michel P.H.
APPLICANT: Guern, Jean R.D.
TITLE OF INVENTION: Method for Promoting the Differentiation
TITLE OF INVENTION: of Plant Cells in Culture
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pollock, Vande Sande & Priddy
STREET: 1990 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-3425
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,481
FILING DATE: 08-AUG-1995
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/01239
FILING DATE: 14-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92 15044
FILING DATE: 14-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Franklin, Eric J.
REGISTRATION NUMBER: 37,134
REFERENCE/DOCKET NUMBER: 0061/00065
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 331-7111
TELEFAX: (202) 293-6229
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-448-481-4

Query Match 66.7%; Score 30; DB 2; Length 40;
Best Local Similarity 83.3%; Pred. No. 42;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 NVPVPM 8
111111
DB 23 NVPVPL 28

RESULT 14
US-08-377-687-41
Sequence 41, Application US/08377687
Patent No. 5538525
GENERAL INFORMATION:
APPLICANT: BROEKERT, WILLEM F.
APPLICANT: CAMMUE, BRUNO P.A.
APPLICANT: OSBORN, RUPERT W.
APPLICANT: REES, SARAH B.
APPLICANT: TERRAS, FRANKY R.G.
APPLICANT: VANDERLEYDEN, JOZER
TITLE OF INVENTION: BIOCIDAL PROTEINS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/377,687
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/002,480
FILING DATE: 04-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:

LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-377-687-41

Search completed: June 13, 2002, 08:54:34
Job time: 319 sec

Query Match 66.7%; Score 30; DB 1; Length 45;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 NVPVPM 8
DB 23 NVPVPL 28

RESULT 15
US-08-777-192-41
Sequence 41, Application US/08777192
Patent No. 5624869
GENERAL INFORMATION:
APPLICANT: BROEKERT, WILLEM F.
APPLICANT: CAMMUE, BRUNO P. A.
APPLICANT: OSBORN, RUPERT W.
APPLICANT: REES, SARAH B.
APPLICANT: TERRAS, FRANKY R. G.
APPLICANT: VANDERLEIJDEN, JOZEF
TITLE OF INVENTION: BIOCIDAL PROTEINS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/777,192
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/002,480
FILING DATE: 04-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: KOKUTIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 99042/SEE. 36525/US/A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
INFORMATION FOR SEQ. ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-777-192-41

Query Match 66.7%; Score 30; DB 2; Length 45;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 NVPVPM 8
DB 23 NVPVPL 28

Fri Jun 14 10:36:51 2002

us-09-710-339-2_copy_468_475.rai

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:55:56 ; Search time 73.31 Seconds

(without alignments)
10.486 Million cell updates/sec

Title: US-09-710-339-2_COPY_468_475

Perfect score: 45

Sequence: 1 DGNVPVPM 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

PIR-71:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	478	2 JK0201	alpha-amy-lase (EC
2	45	100.0	498	2 A48305	alpha-amy-lase (EC
3	45	100.0	499	1 ALAS1	alpha-amy-lase (EC
4	45	100.0	499	1 ALAS3	alpha-amy-lase (EC
5	45	100.0	499	2 B48305	alpha-amy-lase (EC
6	45	100.0	499	2 JS0663	alpha-amy-lase (EC
7	45	100.0	499	2 JT0466	alpha-amy-lase (EC
8	45	100.0	499	2 JN0588	alpha-amy-lase (EC
9	45	100.0	499	2 JN0588	polyketide synthas
10	36	80.0	336	2 AG0511	dtcp-glucose 4,6-d
11	36	80.0	428	2 AG0511	FixC protein limpo
12	36	80.0	446	2 B70776	probable glp2 - My
13	36	80.0	837	1 A31842	endo-1,4-beta-xyla
14	36	80.0	1385	2 S61236	major capsid prote
15	36	80.0	1396	1 VCBE40	auxin-regulated pr
16	35	77.8	612	2 T05750	polyketide synthas
17	35	77.8	8563	2 T30226	polyketide synthas
18	35	77.8	10223	2 T30225	5'-methylthioadeno
19	34	75.6	297	2 E72102	hypothetical prote
20	34	75.6	465	2 T34690	hypothetical prote
21	34	75.6	926	2 T39664	chitin synthase 2
22	34	75.6	1038	2 T13177	soy protein - Trul
23	34	75.6	1390	2 T31353	polypeptide - Arab
24	34	75.6	1974	2 T16703	hypothetical prote
25	33.5	74.4	90	2 H69886	hypothetical prote
26	33	73.3	81	2 JN0058	hypothetical 8.9K
27	33	73.3	125	2 A71024	hypothetical prote
28	33	73.3	261	2 T24532	hypothetical prote
29	33	73.3	346	2 A13587	membrane dipeptida

ALIGNMENTS

30	33	73.3	379	2 A48936	nitrite reductase
31	33	73.3	403	2 C83924	aspartate aminotra
32	33	73.3	430	2 AC3037	hypothetical prote
33	33	73.3	451	2 S30401	hypothetical prote
34	33	73.3	471	2 H98248	nitrate-binding pr
35	33	73.3	484	1 A35282	alpha-amy-lase (EC
36	33	73.3	507	2 T08337	hypothetical prote
37	33	73.3	671	2 A97352	penicillin-binding
38	33	73.3	760	2 AB2225	hypothetical prote
39	33	73.3	870	2 A86567	DNA polymerase I C
40	33	73.3	870	2 F72058	DNA polymerase I C
41	33	73.3	1061	2 T13107	tail tip fiber pro
42	33	73.3	1375	2 T42586	major capsid prote
43	33	73.3	1376	1 VCBED6	major capsid prote
44	33	73.3	1484	2 T29275	hypothetical prote
45	33	71.1	149	2 C72410	actinorhodin polyk

RESULT 1
JK0201
Alpha-amy-lase (EC 3.2.1.1) - Aspergillus oryzae
N:Alternate names: glycogenase; Taka-amy-lase A
C:Species: Aspergillus oryzae
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: JK0201
R:Toda, H.; Kondo, K.; Narita, K.
Proc. Jpn. Acad. 58B, 208-212, 1982
A:Title: The complete amino acid sequence of Taka-amy-lase A.
A:Reference number: JK0201
A:Accession: JK0201
A:Molecule type: protein
A:Residues: 1-478 <TOP>
C:Comment: One atom of calcium per molecule is essential for the activity.
C:Comment: This enzyme is a glycoprotein.
C:Comment: See also PIR:JT0466 and PIR:JS0240.
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amy-lase; alpha-amy-lase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:173-300/Domain: alpha-amy-lase core homology <AMT>
F:197/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 45; DB 2; Length 478;
Best local Similarity 100.0%; Pred. No. 0.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DGNVPVPM 8
Db 447 DGNVPVPM 454

RESULT 2
A48305
Alpha-amy-lase (EC 3.2.1.1) A precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: A48305
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmone, C.L.; Kodama, K.H.; Royer, T.J.
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amy-lase genes from As
A:Reference number: A48305; WUID:90254827
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-498 <KOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amylose core homology <AMT>

Query Match 100.0%; Score 45; DB 2; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DGNVPPM 8
 Db 469 DGNVPPM 476

RESULT 3
 A:AS1
 alpha-amylose (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae
 M:Alternate names: alpha-amylose isozyme II; glycogenase; taka-amylose A
 C:Species: Aspergillus oryzae
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 15-Sep-2000
 C:Accession: S04548; A33214; J50240; A91930; A93767; A10627
 R:Witsell, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 M:1. Microbiol. 3, 3-14, 1989
 A:Title: Three alpha-amylose genes of Aspergillus oryzae exhibit identical intron-exon
 A:Reference number: S04548; MUID:89237897
 A:Accession: S04548
 A:Molecule type: DNA
 A:Residues: 1-499 <MT>
 A:Cross-references: GB:X12725; NID:92430; PIDN:CAA31218.1; PID:9295921
 R:Genes, M.J.; Dove, M.J.; Seligy, V.L.
 Gene 79, 107-117, 1989
 A:Title: Aspergillus oryzae has two nearly identical Taka-amylose genes, each containing
 A:Reference number: J50240; MUID:89378767
 A:Accession: J50240
 A:Molecule type: DNA
 A:Residues: 1-499 <GEN>
 A:Genetics: AMT2
 A:Note: The authors refer to this as isozyme II
 R:Isemura, S.; Ikenaka, T.
 J. Biochem. 74, 1-10, 1973
 A:Reference number: A91930; MUID:74001521
 A:Accession: A91930
 A:Molecule type: protein
 A:Residues: 206-225 <ISE>
 R:Narita, K.
 Proc. Jpn. Acad. 51, 285-290, 1975
 A:Reference number: A93767
 A:Accession: A93767
 A:Molecule type: protein
 A:Residues: 434-443,446-447, 'Q', 449-458, 'GTYV', 459-464,467-468, 'B', 470, 'B', 472-499 <NMR>
 R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
 J. Biochem. 95, 697-702, 1984
 A:Title: Structure and possible catalytic residues of Taka-amylose A.
 A:Reference number: A37454; MUID:84212370
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms
 R:Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkemburg, J.P.; Wilkinson, A.
 submitted to the Brookhaven Protein Data Bank, August 1992
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics: <AMT1>
 A:Gene: amy1
 A:introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Genetics: <AMT2>
 A:Gene: amy2; Amy11
 A:introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; po
 F:1-21/Domain: signal sequence #status predicted <St>
 F:12-499/Product: alpha-amylose 1 #status experimental <MAT>
 F:194-321/Domain: alpha-amylose core homology <AMT>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 45; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DGNVPPM 8
 Db 469 DGNVPPM 476

RESULT 4
 A:AS3
 alpha-amylose (EC 3.2.1.1) 3 precursor - Aspergillus oryzae
 M:Alternate names: alpha-amylose isozyme I; glycogenase; Taka-amylose A
 C:Species: Aspergillus oryzae
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
 C:Accession: S04549; A33215; A44713
 R:Witsell, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 M:1. Microbiol. 3, 3-14, 1989
 A:Title: Three alpha-amylose genes of Aspergillus oryzae exhibit identical intron-exo
 A:Reference number: S04548; MUID:89237897
 A:Accession: S04549
 A:Molecule type: DNA
 A:Residues: 1-499 <MT>
 A:Cross-references: EMBL:X12727; NID:92454; PIDN:CAA31220.1; PID:9295922
 R:Genes, M.J.; Dove, M.J.; Seligy, V.L.
 Gene 79, 107-117, 1989
 A:Title: Aspergillus oryzae has two nearly identical Taka-amylose genes, each contain
 A:Reference number: J50240; MUID:89378767
 A:Accession: A44713
 A:Molecule type: DNA
 A:Residues: 1-499 <GEN>
 A:Note: The authors refer to this as isozyme I
 R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
 J. Biochem. 95, 697-702, 1984
 A:Title: Structure and possible catalytic residues of Taka-amylose A.
 A:Reference number: A37454; MUID:84212370
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics:
 A:Gene: amy3; Amy1
 A:introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me
 F:1-21/Domain: signal sequence #status predicted <St>
 F:12-499/Product: alpha-amylose 3 #status experimental <MAT>
 F:194-321/Domain: alpha-amylose core homology <AMT>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 45; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
 |||||
 DB 469 DGNVPVP 476

RESULT 5
 B48305
 alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori
 C:Species: Aspergillus awamori
 C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
 C:Accession: B48305
 R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;
 Curr. Genet. 17, 203-212, 1990
 A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Asper
 A:Reference number: A48305; MUID:90254827
 A:Accession: B48305
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1499 <KOR>
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Pathway: glycogen/starch degradation
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <Sig>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 45; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
 |||||
 DB 469 DGNVPVP 476

RESULT 6
 JS0663
 alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
 C:Species: Aspergillus sp.
 C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
 C:Accession: JS0663
 R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
 Biosci. Biotechnol. Biochem. 56, 174-179, 1992
 A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shirousanii and its expression
 A:Reference number: JS0663; MUID:92323146
 A:Accession: JS0663
 A:Molecule type: mRNA
 A:Residues: 1499 <SHI>
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <Sig>
 F:22-499/Product: alpha-amylase #status predicted <AMP>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 45; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
 |||||
 DB 469 DGNVPVP 476

RESULT 7
 JT0466
 alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
 N:Alternate names: glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae
 C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
 C:Accession: JT0466
 R:Tada, S.; Imura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
 Agric. Biol. Chem. 53, 593-599, 1989
 A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Asperg
 A:Reference number: JT0466
 A:Accession: JT0466
 A:Molecule type: DNA
 A:Residues: 1499 <TAD>
 C:Comment: See also PIR:JK0201 and PIR:JS0240.
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics:
 A:Insertions: 56/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <Sig>
 F:22-499/Product: alpha-amylase #status predicted <AMY>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 45; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
 |||||
 DB 469 DGNVPVP 476

RESULT 8
 JN0588
 alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
 N:Alternate names: Taka-amylase A
 C:Species: Aspergillus oryzae
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
 C:Accession: JN0588
 R:Tsukagoshi, N.; Furukawa, M.; Nagaba, H.; Kirita, N.; Tsudoi, A.; Udaoka, S.
 Gene 84, 319-327, 1989
 A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for
 A:Reference number: JN0588; MUID:90128276
 A:Accession: JN0588
 A:Molecule type: mRNA
 A:Residues: 1499 <TSU>
 C:Comment: The alpha amylases are encoded by multigene family.
 C:Genetics:
 A:Gene: Taa-G1
 A:Insertions: 56/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <Sig>
 F:22-499/Product: alpha-amylase #status predicted <AMY>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 45; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8

Db 469 DGNVPVP 476
|||||

RESULT 9
T30228
polyketide synthase - Streptomyces hygroscopicus
C:Species: Streptomyces hygroscopicus
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
C:Accession: T30228
R:Apriorio, J.F.; Molnar, I.; Schwecke, T.; Koenig, A.; Haydock, S.F.; Khaw, L.E.; Straub
Gene 169, 9-16, 1996
A:Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces hyg
A:Reference number: 220782; PMID:9618686
A:Accession: T30228
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6260 <PAP>
A:Cross-references: EMBL:X86780; NID:g987088; PID:g987102; PIDN:CAA60462.1
C:Genetics:
A:Gene: rAPC
C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protei
C:Keywords: carrier protein
F:55-451/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F:1536-1607/Domain: acyl carrier protein homology <ACPI>
F:1651-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F:3123-3194/Domain: acyl carrier protein homology <ACP2>
F:3228-3632/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F:3725-3999/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F:5114-5185/Domain: acyl carrier protein homology <ACP3>
F:5242-5638/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F:5759-6044/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F:6135-6206/Domain: acyl carrier protein homology <ACP4>

Query Match 82.2%; Score 37; DB 2; Length 6260;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 7
|||||
Db 4372 DGNMPVP 4378

RESULT 10
AG3564
dcdp-glucose 4,6-dehydratase (EC 4.2.1.46) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AG3564
R:DeVecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mufer, C.; Los, T.; Ivanova,
Mazur, M.; Goldsmun, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AG3564
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-536 <KUR>
A:Cross-references: GB:AE008918; PIDN:AM53682.1; PID:g17984602; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME110440
A:Map position: 11
C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 80.0%; Score 36; DB 2; Length 336;
Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 DGNVPVP 8
|||||

Db 186 DNNLPVP 193

RESULT 11
AG0511
fixC protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain C
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AG0511
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quall, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A:Reference number: AB0502; PMID:11677608
A:Accession: AG0511
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-428 <PAR>
A:Cross-references: GB:AL51382; PIDN:CAD01231.1; PID:g16501360; GSPDB:GN00176
C:Genetics:
A:Gene: STY0087
C:Superfamily: fixC protein

Query Match 80.0%; Score 36; DB 2; Length 428;
Best Local Similarity 75.0%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
|||||
Db 394 DGSAPVP 401

RESULT 12
B70776
probable gln2 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70776
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; PMID:98295987
A:Accession: B70776
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-446 <COL>
A:Cross-references: GB:770692; GB:AL123456; NID:g3261567; PIDN:CAA94645.1; PID:g12619
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: glnA2
C:Superfamily: glutamate--ammonia ligase

Query Match 80.0%; Score 36; DB 2; Length 446;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
|||||
Db 144 DGSVPVP 151

RESULT 13
A31842
endo-1,4-beta-xylanase (EC 3.2.1.8) z precursor - Clostridium thermocellum
N:Alternate names: xylanase Z

C:Species: Clostridium thermocellum
C:Date: 31-Mar-1990 #sequence_revision 11-Apr-1997 #text_change 18-Jun-1999
C:Accession: A31842
R:Grepinet, O.; Chebrou, M.C.; Beguin, P.
J. Bacteriol. 170, 4582-4588, 1988
A:Title: Nucleotide sequence and deletion analysis of the xylanase gene (xynZ) of Clostridium thermocellum
A:Reference number: A31842; MUID:89008072
A:Accession: A31842
A:Molecule type: DNA
A:Residues: 1-837 <GRE>
A:Cross-references: GB:M2624; NID:g144931; PIDN:AAA23286.1; PID:g144932
C:Genetics:
A:Gene: xynZ
C:Function:
A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A:Pathway: xylan degradation
C:Superfamily: Clostridium endo-1,4-beta-xylanase Z; Clostridium cellulase repeat homolog
C:Keywords: duplication; extracellular protein; glycosidase; heat-stable protein; hydrolysis
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-837/Product: endo-1,4-beta-xylanase #status predicted <MAT>
F:326-419/Domain: Clostridium xylanase A repeat homology <CXA>
F:430-453/Domain: Clostridium cellulase repeat homology <CCR1>
F:464-487/Domain: Clostridium cellulase repeat homology <CCR2>
F:548-834/Domain: Streptomyces endo-1,4-beta-xylanase A homology <SXY>
F:645,754/Active site: Glu #status predicted

Query Match 80.0%; Score 36; DB 1; Length 837;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DGNVPVP 7
DB 284 DGNVPVP 290

RESULT 14
S61236
major capsid protein - bovine herpesvirus 1
C:Species: bovine herpesvirus 1
C:Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 26-Aug-1999
C:Accession: S61236
R:Vilek, C.; Benes, V.; Lu, Z.; Kutish, G.F.; Paces, V.; Rock, D.; Letchworth, G.J.; Schmitt, J.; et al.
submitted to the EMBL Data Library, January 1995
A:Description: Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus 1 genome
A:Reference number: S61233
A:Accession: S61236
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1385 <VIC>
A:Cross-references: EMBL:Z48053; NID:g971311; PIDN:CAA8114.1; PID:g971315
C:Superfamily: varicella-zoster virus major capsid protein

Query Match 80.0%; Score 36; DB 2; Length 1385;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 DGNVPVP 8
DB 580 NGNIPVPL 587

RESULT 15
VCB840
major capsid protein - human herpesvirus 3
C:Species: human herpesvirus 3
C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
C:Accession: E27341
R:Davidson, A.J.; Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A:Title: The complete DNA sequence of varicella-zoster virus.
A:Reference number: A27345; MUID:86306657

A:Accession: E27341
A:Molecule type: DNA
A:Residues: 1-1396 <DAV>
A:Cross-references: EMBL:X04370; NID:g959989; PIDN:CAA27923.1; PID:g60029
C:Genetics:
A:Gene: 40
C:Superfamily: varicella-zoster virus major capsid protein
C:Keywords: capsid protein

Query Match 80.0%; Score 36; DB 1; Length 1396;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 DGNVPVP 8
DB 595 NGNIPVPL 602

Search completed: June 13, 2002, 08:55:58
Job time: 378 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:55 ; Search time 38.1 Seconds

(without alignments)
8.130 Million cell updates/sec

Title: US-09-710-339-2_COPY_468_475

Perfect score: 45

Sequence: 1 DGNVPVPM 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	498	1	AMYA_ASPAW
2	45	100.0	499	1	AMYA_ASPOR
3	45	100.0	499	1	AMYA_ASPAW
4	45	100.0	499	1	AMYA_ASPAW
5	36	80.0	364	1	Y174_HUMAN
6	36	80.0	446	1	GLN2_MYCTU
7	36	80.0	837	1	XYNZ_CLOTM
8	36	80.0	1396	1	VCAP_VZVD
9	34	75.6	926	1	CHS2_SCHPO
10	34	75.6	1038	1	SOE_DROME
11	34	75.6	1654	1	PCFB_HUMAN
12	33	73.3	125	1	RT06_MOUSE
13	33	73.3	379	1	NIR_PSESP
14	33	73.3	484	1	AMYA_ASPNG
15	33	73.3	1376	1	VCAP_HSYVJ
16	32	71.1	380	1	GBR_ORYSA
17	32	71.1	528	1	VP40_HSYVU
18	32	71.1	678	1	CMC1_HUMAN
19	32	71.1	683	1	AMOH_ARTGO
20	32	71.1	1345	1	VCAP_HSYVU
21	32	71.1	1345	1	VCAP_HSYVJ
22	31	68.9	124	1	RNP_PIG
23	31	68.9	131	1	Y1L2_STRAU
24	31	68.9	262	1	KK4_BACCT
25	31	68.9	327	1	F1P1_YEAST
26	31	68.9	397	1	TBXT_CHICK
27	31	68.9	514	1	UVS2_XENLA
28	31	68.9	521	1	PHLB_MYCTU
29	31	68.9	680	1	APB2_MOUSE
30	31	68.9	1066	1	HGPC_HAEIN
31	31	68.9	1067	1	HGBB_HAEIN
32	31	68.9	1109	1	MSH7_ARATH
33	31	68.9	1505	1	DPOA_DROME

ALIGNMENTS

RESULT	1	AMYA_ASPAW	STANDARD	PRT	498 AA
AC	002905;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	01-OCT-1996 (Rel. 34, Last annotation update)				
DE	Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase A).				
DE	glucanohydrolase A).				
GN	AMYA.				
OS	Aspergillus awamori.				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
OC	Eurotiaceae; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				
OX	NCBI_TaxID=105351;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=UVR143F.				
RX	MEDLINE=90254827; PubMed=2340591;				
RA	Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,				
RA	Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,				
RA	Berka R.M.;				
RT	"Cloning, characterization, and expression of two alpha-amylase genes				
RT	from Aspergillus niger var. awamori.";				
RT	Curr. Genet. 17:203-212(1990).				
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic				
CC	linkages in oligosaccharides and polysaccharides.				
CC	-1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO				
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: X52755; CNA36966.1; -				
DR	HSSP: P10529; 77RA.				
DR	Interp: IPRO0461; Alpha-amylase.				
DR	Plan: PP00128; alpha-amylase; 1.				
KW	Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;				
KW	Glycoprotein; Signal; Multigene family.				
FT	SIGNAL 1 21				POTENTIAL.
FT	CHAIN 1 22				ALPHA-AMYLASE A.
FT	ACT_SITE 227				BY SIMILARITY.
FT	ACT_SITE 231				BY SIMILARITY.
FT	ACT_SITE 251				BY SIMILARITY.
FT	ACT_SITE 318				BY SIMILARITY.
FT	DISULFID 51				BY SIMILARITY.
FT	DISULFID 171				BY SIMILARITY.
FT	DISULFID 261				BY SIMILARITY.
FT	DISULFID 461				BY SIMILARITY.
FT	CARBOHYD 218				N-LINKED (GLCNAC...) (POTENTIAL).
FT	SEQUENCE 498 AA: 54880 MW: 765851				IBCO1A8A01 CRC64;

34	30	66.7	87	1	DB15_MOUSE
35	30	66.7	87	1	DB15_RAT
36	30	66.7	120	1	MTTP_DAUCA
37	30	66.7	171	1	Y708_TREPA
38	30	66.7	189	1	DYR_CHICK
39	30	66.7	201	1	CYC4_PSEAE
40	30	66.7	210	1	CYC4_PSEAE
41	30	66.7	241	1	UTR4_YEAST
42	30	66.7	268	1	VG65_BPML5
43	30	66.7	300	1	P1MH_ECOLI
44	30	66.7	336	1	E13A_LYCES
45	30	66.7	338	1	IPNS_CEPAC

009035	mus musculu
P56702	retius norv
P27631	daucus caro
O83706	treponema p
P00378	gallus gall
P00106	pseudomonas
Q52369	pseudomonas
P32626	saccharomyc
Q05279	mycobacteri
P08191	escherichia
O01412	lycopersico
P03189	cephalospo

Query Match 100.0%; Score 45; DB 1; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DGNVPEPM 8
 Db 469 DGNVPEPM 476

RESULT 2
 ID AMYA-ASPOR STANDARD: PRT: 499 AA.
 AC P10529; P11763; Q00250;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amyase A precursor (EC 3.2.1.1) (Taka-amyase A) (TAA)
 DE (1,4-alpha-D-glucan glucanohydrolase).
 GN AMY1 AND AMY2 AND AMY3.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5062;
 RN [1]
 RP SEQUENCE FROM N.A..
 RC STRAIN=DSM 63303;
 RA MEDLINE=89237897; PubMed=2785629;
 RA Witsell S., Lachmund A., Wildhardt G., Rutkowski E.;
 RT "Three alpha-amyase genes of Aspergillus oryzae exhibit identical
 RT intron-exon organization.";
 RL Mol. Microbiol. 3:3-14(1989).
 RN [2]
 RP SEQUENCE FROM N.A..
 RX MEDLINE=89378767; PubMed=2789162;
 RA Genes M.J., Dove M.J., Selligy V.L.;
 RT "Aspergillus oryzae has two nearly identical Taka-amyase genes, each
 RT containing eight introns.";
 RL Gene 79:107-117(1989).
 RN [3]
 RP SEQUENCE FROM N.A..
 RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
 RT "Cloning and nucleotide sequence of the genomic Taka-amyase A gene
 RT of Aspergillus oryzae.";
 RL Agric. Biol. Chem. 53:593-599(1989).
 RN [4]
 RP SEQUENCE FROM N.A..
 RX MEDLINE=90128276; PubMed=2612911;
 RA Tsubagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A.,
 RA Uekura S.;
 RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amyase A:
 RT evidence for multiple related genes.";
 RL Gene 84:319-327(1989).
 RN [5]
 RP SEQUENCE OF 22-499.
 RA Toda H., Kondo K., Narita K.;
 RT "The complete amino acid sequence of Taka-amyase A";
 RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
 RN [6]
 RP SEQUENCE OF 206-225.
 RX MEDLINE=74001521; PubMed=4733850;
 RA Isemura S., Ikenaka T.;
 RT "The amino acid sequences of glycopeptides obtained from Taka-amyase
 RT A with trypsin and chymotrypsin.";
 RL J. Biochem. 74:1-10(1973).
 RN [7]
 RP SEQUENCE OF 433-499.
 RA Narita K.;
 RT Proc. Jpn. Acad., B, Phys. Biol. Sci. 51:285-290(1975).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=80227691; PubMed=6156152;

RA Matsura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
 RA Toda H., Narita K., Kakudo M.;
 RT "Molecular structure of taka-amyase A. I. Backbone chain folding at
 RT 3-A resolution.";
 RL J. Biochem. 87:1555-1558(1980).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=84212370; PubMed=6609921;
 RA Matsura Y., Kusunoki M., Harada W., Kakudo M.;
 RT "Structure and possible catalytic residues of Taka-amyase A.";
 RL J. Biochem. 95:697-702(1984).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
 RX MEDLINE=97428212; PubMed=9283074;
 RA Brozowski A.M., Davies G.J.;
 RT "Structure of the Aspergillus oryzae alpha-amyase complexed with the
 RT inhibitor acarbose at 2.0-A resolution.";
 RL Biochemistry 36:10837-10845(1997).
 CC -I- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -I- COPFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -I- MISCELLANEOUS: THE SEQUENCE OF AMY1 AND AMY2 IS SHOWN.
 CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLAASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 DR EMBL: X12725; CAA31218.1; -;
 DR EMBL: X12726; CAA31219.1; -;
 DR EMBL: X12727; CAA31220.1; -;
 DR EMBL: D00434; BAA00336.1; -;
 DR EMBL: M33218; AAA32708.1; -;
 DR PIR: S04548; ALAS1.
 DR PIR: S04549; ALAS3.
 DR PIR: A10627; A10627.
 DR PIR: JK0201; JK0201.
 DR PIR: JS0240; JS0240.
 DR PIR: JT0466; JT0466.
 DR PDB: 2TAA; 15-OCT-89.
 DR PDB: 6TAA; 31-OCT-93.
 DR PDB: 7TAA; 25-NOV-98.
 DR GlycoSuiteDB: P10529;
 DR InterPro: IPR000461; Alpha-amyase.
 DR Pfam: PF00128; alpha-amyase; 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 304 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 FT VARIANT 56 56
 FT VARIANT 172 172
 FT CONFLICT 93 94
 FT CONFLICT 106 106
 FT CONFLICT 184 184
 FT CONFLICT 195 195
 FT CONFLICT 255 255
 FT CONFLICT 291 291
 FT CONFLICT 345 345
 N-LINKED (GLCNAC. . .).
 /FTID-CAR.000125.
 O -> R (IN AMY3).
 F -> L (IN AMY3).
 FT -> DC (IN REF. 5).
 O -> T (IN REF. 5).
 D -> Y (IN REF. 3).
 P -> L (IN REF. 3).
 G -> V (IN REF. 3).
 D -> H (IN REF. 4).
 I -> L (IN REF. 5).

FT CONFLICT 370 370 L -> A (IN REF. 4).
 FT CONFLICT 406 409 MP1Y -> PYI (IN REF. 5).
 FT CONFLICT 448 448 G -> S (IN REF. 5).
 FT CONFLICT 497 497 S -> SD (IN REF. 5 AND 7).

FT HELIX 24 27
 FT TURN 28 29
 FT STRAND 32 35
 FT HELIX 37 40
 FT STRAND 42 42
 FT TURN 43 44
 FT HELIX 53 55
 FT STRAND 61 61
 FT HELIX 63 67
 FT TURN 68 69
 FT HELIX 70 74
 FT TURN 75 77
 FT STRAND 80 83
 FT STRAND 87 89
 FT STRAND 94 94
 FT TURN 95 96
 FT STRAND 97 97
 FT TURN 100 101
 FT STRAND 105 111
 FT TURN 113 115
 FT TURN 118 119
 FT HELIX 120 130
 FT TURN 131 132
 FT STRAND 134 139
 FT STRAND 143 143
 FT STRAND 146 146
 FT HELIX 150 152
 FT TURN 155 156
 FT STRAND 159 159
 FT HELIX 164 166
 FT STRAND 167 167
 FT STRAND 172 172
 FT HELIX 179 184
 FT STRAND 186 188
 FT STRAND 193 194
 FT STRAND 196 197
 FT TURN 199 200
 FT HELIX 202 219
 FT TURN 220 220
 FT STRAND 223 226
 FT TURN 227 228
 FT HELIX 229 231
 FT TURN 234 235
 FT HELIX 236 244
 FT STRAND 247 250
 FT HELIX 257 265
 FT TURN 266 266
 FT STRAND 269 271
 FT HELIX 273 283
 FT TURN 286 287
 FT HELIX 290 303
 FT STRAND 307 309
 FT TURN 311 312
 FT TURN 317 318
 FT HELIX 322 325
 FT HELIX 329 341
 FT STRAND 345 349
 FT TURN 350 351
 FT HELIX 352 354
 FT TURN 355 355
 FT TURN 361 364
 FT HELIX 368 371
 FT TURN 372 372

Query Match 100.0%; Score 45; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DGNVPVP 8

Db 469 DGNVPVP 476

RESULT 3
 ID AMYB_ASPAW STANDARD; PRT; 499 AA.
 AC Q02906;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Alpha-amyase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
 DE glucanohydrolase B).
 GN AMYB.
 OS Aspergillus awamori.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=105351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DKV143F;
 RX MEDLINE=90254827; PubMed=2340591;
 RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
 RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
 RA Berka R.M.;
 RT "Cloning, characterization, and expression of two alpha-amyase genes
 RT from Aspergillus niger var. awamori";
 RL Curr. Genet. 17:203-212(1990).
 CC -I- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -I- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: X52756; CAA36967.1; -.
 DR HSSP: P10529; YTA.
 DR InterPro: IPR000461; Alpha_amyase.
 DR Pfam: PF00128; alpha-amyase; 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 SQ SEQUENCE 499 AA; 54921 MW; 740B96B1BC01A8A CRC64;
 N-LINKED (GLCNAC...) (POTENTIAL).

Query Match 100.0%; Score 45; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DGNVPVP 8
 Db 469 DGNVPVP 476
 RESULT 4
 AMT_ASPSH

```

ID  AMY_ASPSH  STANDARD;  PRT;  499 AA.
AC  P30292;
DT  01-APR-1993 (Rel. 25, Created)
DT  01-APR-1993 (Rel. 25, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE  glucanohydrolase).
GN  AMY.
OS  Aspergillus shirousami.
OC  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC  Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
RX  NCBI_TaxID=5070;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92323146; PubMed=1368777;
RX  Shibuya I., Tamura G., Ishikawa T., Hara S.;
RT  "Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its
RT  expression in Saccharomyces cerevisiae."
RL  Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC  -1 CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC  linkages in oligosaccharides and polysaccharides.
CC  -1 COPOLYMER: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC  -1 SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: D10461; BAA01255.1; -.
DR  PIR: J50663; J50663.
DR  HSSP: P10529; 7TRA.
DR  InterPro: IPR000461; Alpha_amylase.
DR  Pfam: PF00128; alpha-amylase; 1.
KM  Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KM  Glycoprotein; Signal.
FT  SIGNAL 1 21
FT  CHAIN 22 499 ALPHA-AMYLASE.
FT  ACT_SITE 227 227
FT  ACT_SITE 231 231
FT  ACT_SITE 251 251 BY SIMILARITY.
FT  ACT_SITE 318 318 BY SIMILARITY.
FT  DISULFID 51 59 BY SIMILARITY.
FT  DISULFID 171 185 BY SIMILARITY.
FT  DISULFID 261 304 BY SIMILARITY.
FT  DISULFID 461 496 BY SIMILARITY.
FT  CARBOHYD 218 218 N-LINKED (GLCNAC... (BY SIMILARITY).
SQ  SEQUENCE 499 AA; 34652 MW; 1F87AE50DA01C03F CRC64;

```

```

Query Match      100.0%; Score 45; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY  1 DGNVPEPM 8
    |||
Db  469 DGNVPEPM 476

```

```

RESULT 5
Y174_HUMAN STANDARD; PRT; 364 AA.
AC  P53990;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Hypothetical protein KIAA0174.
GN  KIAA0174.
OS  Homo sapiens (Human).

```

```

OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=96281124; PubMed=8724849;
RX  Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;
RT  "Prediction of the coding sequences of unidentified human genes. V.
RT  the coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by
RT  analysis of cDNA clones from human cell line KG-1."
RL  DNA Res. 3:17-24(1996).
CC  -1 SIMILARITY: TO YEAST YNL265C.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: D79996; BAA11491.1; -.
DR  Hypothetical protein.
SQ  SEQUENCE 364 AA; 39750 MW; 0DD3C186A52A4380 CRC64;

```

```

Query Match      80.0%; Score 36; DB 1; Length 364;
Best Local Similarity 75.0%; Pred. No. 9.6;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY  1 DGNVPEPM 8
    |||
Db  225 DGVPEPM 232

```

```

RESULT 6
G1N2_MYCTU STANDARD; PRT; 446 AA.
AC  Q10378;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Probable glutamine synthetase 2 (EC 6.3.1.2) (Glutamate--ammonia
DE  lyase 2).
GN  G1NA2 OR RV2222C OR MT2280 OR MTCY427.03C OR MTCY190.33C.
OS  Mycobacterium tuberculosis.
OC  Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC  Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX  NCBI_TaxID=1773;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=H37RV;
RX  MEDLINE=98295987; PubMed=9634230;
RX  Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RX  Gordon S.V., Eiglmier K., Gas S., Barry C.E. III, Tekala F.,
RX  Baddock K., Basham D., Brown D., Chillingworth T., Connor R.,
RX  Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RX  Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RX  Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RX  Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RX  Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT  "Deciphering the biology of Mycobacterium tuberculosis from the
RT  complete genome sequence."
RL  Nature 393:537-544(1998).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  STRAIN=COC 1551 / Oshkosh;
RX  Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RX  Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RX  Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RX  Delcher A., Utterback T., Weidman J., Khouri H., Gali J., Mikula A.,
RX  Bishal W.;

```

```

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -1- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO
CC HEXAGON (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z70692; CAA94645.1; -.
DR EMBL: AE007073; AAK46565.1; -.
DR HSSP: P06201; ILGR.
DR TIGR: MT2280; -.
DR TruncatList: RV2222c; -.
DR InterPro: IPR001691; GLN_synth.
DR InterPro: IPR001637; GLN_adenyln.
DR Pfam: PF00120; gln-synt; 1.
DR ProDom: PD001057; GLN_adenyln; 1.
DR PROSITE: PS00180; GLN_1; FALSE_NEG.
DR PROSITE: PS00181; GLN_ATP; FALSE_NEG.
DR Ligase; Multigene family; Complete proteome.
SQ SEQUENCE 446 AA; 49607 MW; 86F163FD017829DD CRC64;

Query Match 80.0%; Score 36; DB 1; Length 446;
Best Local Similarity 75.0%; Pred. NO. 12;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 8
DB 144 DGSVPVPV 151

RESULT 7
XNZ_CLOTM STANDARD; PRT; 837 AA.
ID XNZ_CLOTM
AC P10478;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Endo-1,4-beta-xylanase Z precursor (EC 3.2.1.8) (Xylanase Z)
DE (1,4-beta-D-xylan xylanohydrolase Z).
GN XNZ.
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1515;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIB 10682;
RC MEDLINE=89008072; PubMed=3139632;
RA Grepinet O., Chebrou M.-C., Beguin P.;
RT "Nucleotide sequence and deletion analysis of the xylanase gene
RT (xynZ) of Clostridium thermocellum.";
RL J. Bacteriol. 170:4582-4588(1988).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS) OF 515-837.
RC STRAIN=NCIB 10682;
RC MEDLINE=95393242; PubMed=7664125;
RA Dominguez R., Souchon H., Spillatelli S., Dauter Z., Wilson K.S.,
RA Chauvaud S., Beguin P., Alzari P.M.;
RT "A common protein fold and similar active site in two distinct
RT families of beta-glycanases.";
RL Nat. Struct. Biol. 2:569-576(1995).

```

```

CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic
CC linkages in xyans.
CC -1- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS
CC WELL AS IN OTHER C.THERMOCCELLUM CELLULOSE ENZYMS. THIS DOMAIN
CC MAY FUNCTION AS THE BINDING LIGAND FOR THE 5L COMPONENT.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY F (FAMILY 10 OF GLYCOSYL
CC HYDROLASES).
CC -1- SIMILARITY: CONTAINS 1 XNZ-TYPE CELLULOSE-BINDING DOMAIN (CBD).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M22624; AAA3286.1; -.
DR PIR: A31842; A31842.
DR PDB: 1XYZ; 29-JAN-96.
DR InterPro: IPR002105; Dockerin_1.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR00379; Est_1ip_thioest_actsite.
DR InterPro: IPR001000; Glyco_hydro_10.
DR Pfam: PF00404; Dockerin_1; 2.
DR Pfam: PF00331; Glyco_hydro_10; 1.
DR PRINTS: PR00134; GLHYDRLASE10.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_2.
DR PROSITE: PS00448; CLOS_CELLULOSE_RPT_2.
DR PROSITE: PS00591; GLYCOSYL_HYDROL_F10; 1.
DR Xylan degradation; Hydrolase; Glycosidase; Repeat; Signal;
KM 3d-structure.
FT SIGNAL 1 28
FT CHAIN 29 837
FT ACT_SITE 645 645
FT ACT_SITE 754 754
FT DOMAIN 328 416
FT DOMAIN 430 487
FT REPEAT 430 453
FT REPEAT 464 487
FT REPEAT 783 789
FT DISULFID 783 789
SQ SEQUENCE 837 AA; 92262 MW; DD4C29F04D12B6CD CRC64;

Query Match 80.0%; Score 36; DB 1; Length 837;
Best Local Similarity 85.7%; Pred. NO. 24;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVP 7
DB 284 DGNVPVP 290

RESULT 8
VCAE_VZVD STANDARD; PRT; 1396 AA.
ID VCAE_VZVD
AC P09245;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major capsid protein (MCP).
GN 40.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=86306657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus.";
RL J. Gen. Virol. 67:1759-1816(1986).
CC -1- FUNCTION: MAJOR PROTEIN OF THE ICOSAEDRAL CAPSID.

```

```

CC -! SIMILARITY: BELONGS TO THE HERPESVIRUSES MAJOR CAPSID PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X0A370; CAA27923.1; -
DR PIR: E27341; VCBR40.
DR InterPro: IPR000912; Herpes_MCP.
DR Pfam: PF03122; Herpes_MCP.1.
DR PRINTS: PR00235; HSVCAPSIDMCP.
DR Coat protein.
DR SEQUENCE 1396 AA; 154979 MW; E509272DB077242F CRC64;

Query Match 80.0%; Score 36; DB 1; Length 1396;
Best Local Similarity 62.5%; Pred. No. 42;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNVPVPM 8
Db 595 NCNIPVPL 602

RESULT 9
CHS2_SCHPO STANDARD; PRT; 926 AA.
ID CHS2_SCHPO
AC 074756;
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitin synthase 2 (Ec 2.4.1.16) (Chitin-UDP acetyl-D-glucosaminyl
DE transferase 2).
DE CHS2 OR SPBC1709.01 OR SPBC1734.17.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Lyne M., Rejandream M.A., Barrell B.G., Rieger M.;
RU Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: PLAYS A MAJOR ROLE IN CELL WALL BIOGENESIS.
CC -! CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + [(1,4)-(N-acetyl-
CC beta-D-glucosaminyl)](N) = UDP + [(1,4)-(N-acetyl-D-
CC glucosaminyl)](N+1).
CC -! SUBCELLULAR LOCATION: Plasma membrane-bound.
CC -! SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL031856; CAA21311.1; -
DR EMBL: AL031852; CAA21240.1; -
DR InterPro: IPR002923; Chitin_synth.
DR Pfam: PF01644; Chitin_synth.1.
DR ProDom: PD002998; Chitin_synth.1.
KW Transferrase; Glycosyltransferase; Transmembrane; Cell wall;
KW Multigene family.
SQ SEQUENCE 926 AA; 105578 MW; 7B880751AD53115E CRC64;

```

```

Query Match 75.6%; Score 34; DB 1; Length 926;
Best Local Similarity 75.0%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DGNVPVPM 8
Db 357 DGNVPMOM 364

RESULT 10
SOG_DROME STANDARD; PRT; 1038 AA.
ID SOG_DROME
AC Q24025; Q9VXS7;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Dorsal-ventral patterning protein sog (short gastrulation protein).
DE SOG OR CG9224.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95047398; PubMed=7958919;
RA Francois V., Solloway M., O'Neill J.W., Emery J., Bier E.;
RT "Dorsal-ventral patterning of the Drosophila embryo depends on a
RT putative negative growth factor encoded by the short gastrulation
RT gene."
RL Genes Dev. 8:2602-2616(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Cocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunn P.,
RA Durbin K., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Gloeck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.D., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Mel M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasco P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maltel B., McIntosh T.C., Meled M.P., Merverson D.,
RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C.C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodedge T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."

```



```

RL Science 287;2185-2195(2000).
CC -1- FUNCTION: PUTATIVE NEGATIVE GROWTH FACTOR; ANTAGONIST OF DPP, A
CC PROTEIN INVOLVED IN PATTERNING THE DORSAL REGION AND IN THE
CC DEVELOPMENT OF THE NEUROECTODERM.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: ABOUT THE DORSAL DPP-EXpressING CELLS IN A
CC LATERAL STRIPE 14-16 CELLS WIDE. LATER IN EMBRYOGENESIS IT IS
CC EXPRESSED IN NEUROECTODERM AND IN THE ENDODERM SPACED ALONG THE
CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING GUT.
CC -1- DEVELOPMENTAL STAGE: EMBRYOGENESIS.
CC -1- SIMILARITY: BELONGS TO THE CHORDIN FAMILY.
CC -1- SIMILARITY: CONTAINS 4 VMFC DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL; U18774; AAA89117.1; -.
DR EMBL; AE003499; AAF48481.1; -.
DR EMBL; F8900003463; sog.
DR InterPro: IPR001007; VMFC.
DR Pfam; PF00093; vwc; 4.
DR SMART; SM00214; vwc; 3.
DR PROSITE; PS01208; VMFC; 2.
DR Transmembrane; Developmental protein; Repeat; Growth factor;
KW Growth regulation; Signal-anchor.
FT DOMAIN 1 53 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 54 74 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT 1 53 (POTENTIAL).
FT DOMAIN 75 1038 EXTRACELLULAR (POTENTIAL).
FT 100 175 VMFC 1.
FT REPEAT 421 522 SR1.
FT REPEAT 592 668 SR2.
FT REPEAT 677 754 SR3.
FT DOMAIN 742 804 VMFC 2.
FT DOMAIN 830 899 VMFC 3.
FT DOMAIN 939 1020 VMFC 4.
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 520 520 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 666 666 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 821 821 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1038 AA; 115514 MW; B0E833AFD79A9037 CRC64;

```

```

Query Match 75.6%; Score 34; DB 1; Length 1038;
Best Local Similarity 85.7%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 DGNVPPV 7
   1 11111
Db 723 DNNVPPV 729

```

```

RESULT 11
PCFB_HUMAN STANDARD; PRT; 1654 AA.
AC 094913; 043671;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pre-mRNA cleavage complex II protein Pcf11 (Fragment).
GN Pcf11 OR KIA0824.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OK NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE OF 11-1654 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9156230; PubMed=10048485;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:355-364(1998).
RN [2]
RP SEQUENCE OF 1-881 FROM N.A.
RA Rouillard J.M., Lacroste F.;
RT "Cloning and sequencing of an Homo sapiens cDNA homologous to the
RT yeast Pcf11 gene.";
RN [3]
RP MEDLINE=20514154; PubMed=11060040;
RX de Vries H., Rueggeberger U., Huebner W., Friedlein A., Langen H.,
RA Keller W.;
RT "Human pre-mRNA cleavage factor II(m) contains homologs of yeast
RT proteins and bridges two other cleavage factors.";
RL EMBL J. 19:5895-5904(2000).
CC -1- FUNCTION: COMPONENT OF PRE-MRNA CLEAVAGE COMPLEX II.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- CAUTION: REF 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 811.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL; AB020631; BAA74847.1; -.
DR EMBL; AF046935; AAC03107.1; ALT_FRAME.
KW mRNA processing; Nuclear protein.
FT NON_TER 1 1
FT DOMAIN 427 535 LYS-RICH.
FT DOMAIN 802 1179 GLY-RICH.
FT DOMAIN 38 41 POLY-ARG.
FT DOMAIN 898 901 POLY-GLY.
FT DOMAIN 941 944 POLY-GLY.
FT DOMAIN 1403 1406 POLY-GLU.
FT DOMAIN 1555 1559 POLY-GLU.
SQ SEQUENCE 1654 AA; 183980 MW; BE29D6E164117BB6 CRC64;

```

```

Query Match 75.6%; Score 34; DB 1; Length 1654;
Best Local Similarity 71.4%; Pred. No. 1,2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 GNVPPVM 8
   1 11111
Db 1303 GNPVPPM 1309

```

```

RESULT 12
RT06_MOUSE STANDARD; PRT; 125 AA.
AC p58064;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Mitochondrial 28S ribosomal protein S6 (MRP-S6).
GN MRPS6 OR RPN56.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OK NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Suzuki T., Terasaki M., Takemoto C., Hanada T., Wada A., Ueda T.,
RA Watanabe K.;
RT "Mammalian mitochondrial ribosome: structural and functional
RT compensation for deficit of RNA with protein components.";
RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: BELONGS TO THE S6P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB049943; BAB40996.1; ALT_INIT.
CC InterPro: IPR000529; Ribosomal_S6.
CC PROSITE: PS01048; RIBOSOMAL_S6: FALSE_NEG.
CC Ribosomal protein; Mitochondrion
CC SEQUENCE 125 AA; 14308 MW; 8140B0BDFDD9025 CRC64;

Query Match
Best Local Similarity 73.3%; Score 33; DB 1; Length 125;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
DB 106 DGIIVPVL 113

RESULT 13
NIR_PSESP
ID NIR_PSESP STANDARD; PRT; 379 AA.
AC 001537;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Copper-containing nitrite reductase precursor (EC 1.7.99.3) (Cu-NIR).
GN NIRU
OS Pseudomonas sp. (strain G-179).
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93175864; PubMed=8439151;
RA Ye R.W., Fries M.R., Bezborodnikov S.G., Averill B.A., Tiedje J.M.;
RT "Characterization of the structural gene encoding a copper-containing
RT nitrite reductase and homology of this gene to DNA of other
RT denitrifiers.";
RL Appl. Environ. Microbiol. 59:250-254(1993).
CC -1- CATALYTIC ACTIVITY: 2 nitric oxide + 2 H(2)O + acceptor = 2
CC nitrite + reduced acceptor.
CC -1- COFACTOR: CONTAINS ONE ATOM OF TYPE I COPPER AND ONE ATOM OF TYPE
CC II COPPER. PDB: TYPE I COPPER IS BOUND WITHIN A SINGLE MONOMER,
CC WHILE TYPE 2 COPPER IS HELD BY RESIDUES FROM EACH OF TWO MONOMERS
CC OF THE TRIMER. NITRITE IS BOUND TO THE TYPE 2 COPPER SITE.
CC PSEUDOCUPROIN IS THE PHYSIOLOGICAL ELECTRON DONOR FOR THE CU-NIR IN
CC VITRO (BY SIMILARITY).
CC -1- PATHWAY: NITRATE ASSIMILATION (DENITRIFICATION PATHWAY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- DOMAIN: THE TYPE I COPPER SITE IN NIR PLAYS A CRUCIAL ROLE FOR
CC ELECTRON TRANSFER FROM PSEUDOCUPROIN TO THE TYPE II COPPER SITE
CC OF NIR, WHICH COMPRISES THE CATALYTIC CENTER OF NIR FOR THE
CC REDUCTION OF NITRITE.
CC -1- SIMILARITY: CONTAINS 2 PLASTOCYANIN-LIKE DOMAINS.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-12 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M97294; AAC79132.1; -.
CC PIR: A48936; A48936.
CC HSSP: P25006; INTF.
CC InterPro: IPR001117; Cu-oxidase.
CC InterPro: IPR001287; Cu-nit rdcase.
CC Pfam: PF00394; Cu-oxidase; 2.
CC PRINTS: PR00695; CUNO2RDXASE.
CC Oxidoreductase; Copper; Metal-binding; Flavoprotein; PDB;
CC Nitrate assimilation; Repeat; Periplasmic; Signal.
CC SIGNAL 1 32
CC CHAIN 33 379 COPPER-CONTAINING NITRITE REDUCTASE.
CC DOMAIN 33 214 PLASTOCYANIN-LIKE 1.
CC METAL 215 379 COPPER (TYPE 1) (BY SIMILARITY).
CC METAL 134 134 COPPER (TYPE 1) (BY SIMILARITY).
CC METAL 139 139 COPPER (TYPE 2) (BY SIMILARITY).
CC METAL 174 174 COPPER (TYPE 2) (BY SIMILARITY).
CC METAL 175 175 COPPER (TYPE 1) (BY SIMILARITY).
CC METAL 184 184 COPPER (TYPE 1) (BY SIMILARITY).
CC METAL 189 189 COPPER (TYPE 1) (BY SIMILARITY).
CC METAL 345 345 COPPER (TYPE 2) (BY SIMILARITY).
CC SEQUENCE 379 AA; 40694 MW; 090A3CBF662F62F CRC64;

Query Match
Best Local Similarity 73.3%; Score 33; DB 1; Length 379;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DGNVPVPM 8
DB 104 DGSVPGPM 111

RESULT 14
AMVA_ASPNG
ID AMVA_ASPNG STANDARD; PRT; 484 AA.
AC P56271;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Acid alpha-amylose (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
OS Aspergillus niger.
CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
CC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
CC NCBI_TaxID=5061;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=91002514; PubMed=2207069;
RA Boel E., Brady L., Brzozowski A.M., Derewenda Z., Dodson G.G.,
RA Jensen V.J., Petersen S.B., Swift H., Thim L., Woldike H.F.;
RT "2.1 Å resolution of two enzymes from Aspergillus niger:
RT alpha-glucan binding in alpha-amyloses; an x-ray diffraction study at
RT Blochemistry 29:6244-6249(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC PDB: 2AAA; 15-JUL-93.
CC InterPro: IPR000461; Alpha-amylose.
CC Pfam: PF00128; alpha-amylose; 1.
CC Hylam; PF00128; Glycosidase; Carbohydrate metabolism; Calcium;
CC Glycoprotein; 3D-structure.
CC ACT_SITE 206 206
CC ACT_SITE 210 210
CC ACT_SITE 230 230
CC ACT_SITE 297 297
CC DISULFID 30 38

```

```

FT DISULFID 150 164
FT DISULFID 240 283
FT DISULFID 440 475
FT CARBOHYD 24 24
FT CARBOHYD 157 157
FT CARBOHYD 197 197
SQ SEQUENCE 484 AA; 52935 MW; 04D596E3468056D CRC64;
      N-LINKED (GLCNAC. . .) (POTENTIAL).
      N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

Query Match      73.3%; Score 33; DB 1; Length 484;
Best Local Similarity 71.4%; Pred. No. 51;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 2 GNAVPPM 8
DB 449 GDIPVPM 455

```

```

RESULT 15
VCAP_HSVR
ID VCAP_HSVR STANDARD; PRT; 1376 AA.
AC P28920;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major capsid protein (MCP) (Capsid protein VP5).
GN 42.
OS Equine herpesvirus type 1 (strain Abap) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92295566; Pubmed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1.";
RL Virology 189:304-316(1992).
CC -1- FUNCTION: MAJOR PROTEIN OF THE ICOSAEDRAL CAPSID.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES MAJOR CAPSID PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M86654; AAB02478.1; -
DR PIR: H36799; VCBED6.
DR InterPro: IPR000912; Herpes_MCP.
DR Pfam: PF03122; Herpes_MCP.1.
DR PRINTS: PR00235; HSVCAPSIDMCP.
KW Coat protein.
SQ SEQUENCE 1376 AA; 152182 MW; C3E866EFAE80AEDB CRC64;

```

```

Query Match      73.3%; Score 33; DB 1; Length 1376;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
OY 1 DGNVPVP 8
DB 578 NGNIPPL 585

```

Search completed: June 13, 2002, 09:11:57
Job time: 1037 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:53:30 ; Search time 158.68 Seconds
(without alignments)
5.600 Million cell updates/sec

Title: US-09-710-339-2_COPY_468_475

Perfect score: 45

Sequence: 1 DGNVPVPM 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: A_Geneseq_032802.*
2: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT:*
23: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	478	15	AAK46065
2	45	100.0	478	16	AAK46065
3	45	100.0	478	16	AAK46065
4	45	100.0	478	16	AAK46065
5	45	100.0	478	16	AAK46065
6	45	100.0	478	16	AAK46065
7	45	100.0	478	16	AAK46065
8	45	100.0	478	16	AAK46065
9	45	100.0	478	16	AAK46065
10	45	100.0	478	16	AAK46065
11	45	100.0	478	16	AAK46065

12	35	77.8	64	14	AAK30639	Epitope of human C
13	34	75.6	380	22	ABG19748	Novel human diagno
14	34	75.6	585	22	ABG67737	Drosophila melanog
15	34	75.6	759	22	ABG07769	Novel human diagno
16	34	75.6	759	22	ABG08475	Novel human diagno
17	34	75.6	1038	22	ABG64317	Drosophila melanog
18	33	73.3	57	22	ABG40405	Peptide #7911 enco
19	33	73.3	57	22	ABG24772	Human brain expres
20	33	73.3	57	22	AAW61213	Protein #6771 enco
21	33	73.3	57	22	AAW73932	Human bone marrow
22	33	73.3	57	22	AAW34111	Peptide #8148 enco
23	33	73.3	77	22	AAW42131	Proionbacterium
24	33	73.3	266	22	AAW2460	C glutamincum prote
25	33	73.3	266	22	AAW76596	Corynebacterium g1
26	33	73.3	870	20	AAW35233	C. pneumoniae prot
27	32	71.1	14	19	AAW78535	SH2 domain binding
28	32	71.1	85	22	ABG15346	Human nervous syst
29	32	71.1	110	22	AAW60846	Proionbacterium
30	32	71.1	100	21	AAW36031	zeq may's protein f
31	32	71.1	105	21	AAW42647	Human ORF2411
32	32	71.1	230	18	AAW37694	Amino acid sequenc
33	32	71.1	329	22	AAW39908	Human polypeptide
34	32	71.1	437	20	AAW9655	Human adenosine nu
35	32	71.1	678	20	AAW96554	Human adenosine nu
36	32	71.1	678	22	AAW94292	Human protein sequ
37	32	71.1	680	22	AAW28170	Novel human secret
38	32	71.1	680	22	AAW39247	Human polypeptide
39	32	71.1	684	22	AAW43653	Human polypeptide
40	32	71.1	713	22	AAW41033	Human polypeptide
41	32	71.1	1011	22	ABW57936	Drosophila melanog
42	31	68.9	74	22	AAW55082	Proionbacterium
43	31	68.9	86	21	AAW37863	Neisserial conserv
44	31	68.9	194	22	ABG03338	Novel human diagno
45	31	68.9	201	21	ABW58589	N. meningitidis am

ALIGNMENTS

RESULT 1
ID AAK46065 standard; protein: 478 AA.
XX AAK46065;
AC
XX
XX 18-JUL-1994 (first entry)
DE Mutant alpha-amylase.
XX
XX Methionine substitution; stability; activity; detergent;
KW dishwashing agents; liquifaction agents.
XX
XX Aspergillus oryzae.
OS
XX
XX W09402597-A.
PN
XX 03-FEB-1994.
PD
XX
XX 06-JUL-1993; 93WO-DK00230.
PF
XX 23-JUL-1992; 92DK-0000946.
PR 16-DEC-1992; 92DK-0001503.
PR 15-MAR-1993; 93DK-0000292.
XX
XX (NOVO) NOVO-NORDISK AS.
PA
XX Bisgard-frantzen H, Svendsen A;
XX WPI. 1994-048855/06.
DR
XX Mutant alpha-amylase from Bacillus species comprising a
PT methionine substitution - with improved stability and activity at
PT low pH, for use in detergents, dishwashing agents and

PT liquefaction agents
XX
PS Claim 1; Page 7; 20pp; English.
XX
CC The sequence of that of the *Aspergillus oryzae* alpha amylase, sold
CC commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can
CC be mutated by substitution of one or more of its methionine residues
CC for any amino acid other than cysteine. The mutant alpha-amylase
CC exhibits a better activity level and better stability in the
CC presence of oxidizing agents than previous mutant alpha amylases,
CC and improved thermostability at moderately low pH. The enzyme can
CC be used as an additive for detergents, dishwashing agents and
CC liquidation agents.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 45; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNVPVPM 8
DB 448 dgnvpvpm 455

RESULT 2
ID AAR79025 standard; protein: 478 AA.
XX AAR79025;
XX
XX 22-MAR-1996 (first entry)
XX
DE Mature taka-amylase A.
XX
KM Wild type; neopullulanase; B. steaerothermophilus; mutant;
KM Food industry; modification; hydrophobicity; replacement
KM Insertion; deletion.
XX
OS *Aspergillus oryzae*.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 30..38
FT Disulfide-bond 150..164
FT Disulfide-bond 240..283
FT Disulfide-bond 439..474
XX
XX JP07177891-A.
XX
PD 18-JUL-1995.
XX
PF 31-OCT-1994; 94JP-0288658.
XX
PR 12-NOV-1993; 93JP-0306096.
XX
PA (EZAKI) EZAKI GLICO CO.
XX (NIDE) NEC CORP.
XX
DR WPI; 1995-279919/37.
XX
PT Modifying a transferase by enhancing hydrophobicity of a selected
PT site - increases transfer activity, also new mutant
PT neo-pullulanase(s)
XX
PS Disclosure; Page 10-11; 18pp; English.
XX
CC This sequence represents the mature form of taka-amylase from *A. oryzae*.
CC This sequence was used in a method for the generation of mutant
CC pullulanases for use in the food industry (see also AAR79026-28). The
CC wild type pullulanase enzyme was modified by the method of the
CC invention for enhancing the hydrophobicity of a selected site of the
CC pullulanase. The method comprises replacement of a group in the

CC selected site with a hydrophobic group, replacement of an amino acid
CC with a hydrophobic amino acid, and/or insertion or deletion of a
CC hydrophobic amino acid from the selected site. The method was used
CC to produce neopullulanases Y377F, S422V and M375L.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 45; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNVPVPM 8
DB 447 dgnvpvpm 454

RESULT 3
ID AAR72450 standard; protein: 478 AA.
XX AAR72450;
XX
XX 01-DEC-1995 (first entry)
XX
DE *Aspergillus oryzae* alpha amylase (mature protein).
XX
KM Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
KM washing; Bacillus licheniformis; Bacillus amyloliquefaciens;
KM Bacillus steaerothermophilus; dyeing; bleaching; scouring; textile;
KM thermostable.
XX
OS *Aspergillus oryzae*.
XX
XX W09510603-A.
XX
XX 20-APR-1995.
XX
PF 05-OCT-1994; 94WO-DK00370.
XX
PR 08-OCT-1993; 93DK-0001133.
PR 02-FEB-1994; 94DK-0000140.
XX
XX (NOVO) NOVO-NORDISK AS.
XX
XX Svendsen A, Thellersen M;
XX Van der zee P, Bisgaardfrantzen H, Borchert T;
XX WPI; 1995-161790/21.
XX
DR
XX
PT New Bacillus derived alpha-amylase variants - having amino acid
PT modifications to improve washing and/or dishwashing performance
XX
XX
PS Disclosure; Page 75-76; 105pp; English.
XX
XX Variant alpha amylase enzymes which have improved washing and/or
XX as detergent additives. The enzymes have one or more amino acid
XX residues added, deleted or substituted. The variants can also be
XX used for textile desizing prior to scouring, bleaching and dyeing.
XX The variants have improved thermostability, acid/alkaline stability;
XX low temperature optimum; pH optimum; higher hydrolysis velocity and
XX improved tolerance to other composition constituents, e.g. oxidation
XX agents.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 45; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DGNVPVPM 8
DB 447 dgnvpvpm 454

DB 448 dguvpyvm 455

RESULT 4

ID AAR78270 standard; Protein: 478 AA.

XX AAR78270;

AC AAR78270;

XX 17-JAN-1996 (first entry)

DE Aspergillus oryzae alpha amylase (mature protein).

XX Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;

XX starch; thermostable; methionine; Bacillus licheniformis;

KW Bacillus amyloliquefaciens; Bacillus stearothermophilus;

XX Aspergillus oryzae.

XX Aspergillus oryzae.

XX WO9521247-A1.

XX 10-AUG-1995.

XX 05-OCT-1994; 94WO-DK00371.

XX 02-FEB-1994; 94DK-0000141.

XX (NOVO) NOVO-NORDISK AS.

XX Marcher D, Nilsson TE, Pedersen HH, Toft AH;

DR WPI: 1995-28367/37.

XX Use of an oxidation stable alpha-amylase - for simultaneous desizing

PT and bleaching or scouring of fabrics contg. starch or starch derivs.

XX Disclosure: Page 25-26; 37pp; English.

XX Oxidation stable alpha amylases can be used for the simultaneous

CC desizing and bleaching or scouring of a fabric comprising starch or

CC starch derivatives. They exhibit a better heat stability,

CC especially in the presence of oxidising agents. They are obtained

CC from a parent alpha amylase by replacing one or more methionine

CC residues with any amino acid different from Cys or Met, preferably

CC Leu, Thr, Ala, Gly, Ser, Ile or Asp. The parent alpha amylase is

CC pref. derived from a Bacillus species, although alpha amylases of

CC fungal origin can also be used. This sequence is the wild type

CC (unmodified) alpha amylase of Aspergillus oryzae.

XX Sequence 478 AA;

SQ

Query Match 100.0%; Score 45; DB 16; Length 478;

Best Local Similarity 100.0%; Pred. No. 3.2;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPYPM 8

DB 448 dguvpyvm 455

RESULT 5

AAW14500 standard; protein: 478 AA.

XX AAW14500;

XX 04-JUN-1997 (first entry)

DE Aspergillus oryzae alpha-amylase (mature protein).

XX alpha-amylase; mature; Bacillus licheniformis; termamyl; fungamyl;

KW Aspergillus oryzae; Bacillus amyloliquefaciens; altered property;

KW calcium dependency; substrate binding; stability; pH optimum;

KW thermostability; cleavage; oligosaccharide substrate; dishwashing;

KW washing; detergent additive; fabric desizing; starch liquefaction;

KW sweetener; ethanol production; variant.

XX Aspergillus oryzae.

XX Key Location/Qualifiers

XX Misc-difference 13..45

FT /label= loop 1 modification region

FT /note= "at least one amino acid residue of a parent

FT alpha-amylase (used as a template for a variant)

FT corresponding to 7-23 of AAW14499 is deleted or

FT replaced with a fragment corresponding to

FT this fragment; claim 33"

Misc-difference 14..40

FT /label= loop 1 modification region

FT /note= "preferred region where at least one amino acid

FT residue of a parent alpha-amylase (used as a

FT template for a variant) corresponding to 8-18

FT of AAW14499 is deleted or replaced with a

FT fragment corresponding to this fragment; claim

FT 35"

Misc-difference 28..42

FT /label= loop 1 modification region

FT /note= "at least one amino acid residue of a parent

FT alpha-amylase (used as a template for a variant)

FT corresponding to 12-19 of AAW14499 is deleted or

FT replaced with a fragment corresponding to

FT this fragment; claim 30"

Misc-difference 32..38

FT /label= loop 1 modification region

FT /note= "preferred region where at least one amino acid

FT residue of a parent alpha-amylase (used as a

FT template for a variant) corresponding to 14-15

FT of AAW14499 is deleted or replaced with a

FT fragment corresponding to this fragment; claim

FT 32"

Misc-difference 66..84

FT /label= loop 2 modification region

FT /note= "at least one amino acid residue of a parent

FT alpha-amylase (used as a template for a variant)

FT corresponding to 44-57 of AAW14499 is deleted or

FT replaced with a fragment corresponding to

FT this fragment; claim 18"

Misc-difference 70..78

FT /label= loop 2 modification region

FT /note= "preferred region where at least one amino acid

FT residue of a parent alpha-amylase (used as a

FT template for a variant) corresponding to 48-51

FT of AAW14499 is deleted or replaced with a

FT fragment corresponding to this fragment; claim

FT 20"

Misc-difference 98..210

FT /label= loop 3 modification region

FT /note= "at least one amino acid residue of a parent

FT alpha-amylase (used as a template for a variant)

FT corresponding to 117-185 of AAW14499 is deleted

FT or replaced with a fragment corresponding to

FT this fragment; claim 24"

Misc-difference 102..206

FT /label= loop 3 modification region

FT /note= "preferred region where at least one amino acid

FT residue of a parent alpha-amylase (used as a

FT template for a variant) corresponding to 121-181

FT of AAW14499 is deleted or replaced with a

FT fragment corresponding to this fragment; claim

FT 26"

Misc-difference 121..181

FT /note= "at least one amino acid residue of a parent

FT alpha-amylase (used as a template for a variant)

FT corresponding to this fragment is deleted or

FT replaced with a fragment corresponding to

FT 102-206 of AAW14499; claim 41"

Misc-difference 121..174

FT /note= "preferred region where at least one amino acid

RESULT 7
ABG19750
ID ABG19750 standard; Protein; 314 AA.
XX
AC ABG19750;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19741.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS83937.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 50109; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 314 AA;

Query Match 80.0%; Score 36; DB 22; Length 314;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 DGNVPVPM 8
111111
Db 112 dgtvmpm 119

RESULT 8
ABG19751
ID ABG19751 standard; Protein; 384 AA.
XX
AC ABG19751;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19742.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS83938.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 50110; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 384 AA;

Query Match 80.0%; Score 36; DB 22; Length 384;
Best Local Similarity 75.0%; Pred. No. 1,1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 DGNVPVPM 8
111111
Db 233 dgtvmpm 240
RESULT 9

XX	AAW37905	AAW37905 standard; Protein; 446 AA.	
XX	AAW37905		
AC	AAW37905		
DT	11-SEP-1998	(first entry)	
XX			
DE	Mycobacterium tuberculosis glutamine synthetase II.		
XX			
KW	N-terminal: extracellular protein; abundant; tuberculosis; TB;		
KW	pathogen; disease; treatment; mycobacterial; drug resistant;		
KW	glutamine synthetase II.		
XX			
OS	Mycobacterium tuberculosis.		
XX	W09814613-A1.		
PN			
PD	09-APR-1996.		
XX			
PF	29-SEP-1997; 97WO-US17756.		
XX			
PR	02-OCT-1996; 96US-0724814.		
XX			
PA	(REGC) UNIV CALIFORNIA.		
XX			
PI	Harth G, Horwitz MA;		
XX			
DR	WPI; 1998-240107/21.		
DR	N-PSDB; AAW29071.		
PT	Treatment of infections caused by pathogenic organisms - by		
PT	interfering with the functional activity of an extracellular enzyme		
PT	identified as necessary for the growth or survival of the organism		
XX			
PS	Example 16; Page 31-32; 88pp; English.		
XX			
CC	The sequence is that of glutamine synthetase II which was used as		
CC	part of a method for the treatment of mammalian disease conditions		
CC	associated with infection by a pathogenic organism (PO) comprising:		
CC	(a) identifying at least one extracellular enzyme necessary for the		
CC	growth or survival of the PO; and (b) interfering with the functional		
CC	activity of the extracellular enzyme to inhibit the growth or survival		
CC	of the PO. The method can be used to treat diseases caused by POS such		
CC	as bacteria, protozoa, viruses and fungi. In particular they can		
CC	be used for the treatment of diseases caused by intracellular		
CC	mycobacterial pathogens including M. tuberculosis, M. bovis,		
CC	M. avium, M. kansasii, M. fortuitum, M. chelonae, M. marinum,		
CC	M. scrofulaceum, M. leprae, M. africanum, M. ulcerans and		
CC	M. microti. The methods may be particularly useful for treating		
CC	drug resistant mycobacterium. Using the methods it is possible		
CC	to treat disease conditions associated with sequestered		
CC	intracellular PO without harming the infected host.		
XX			
SO	Sequence 446 AA:		
QY	1 DGNVPVP 8		
	: :		
DB	144 dgsyvpvp 151		
RESULT	10		
ABB62169			
ID	ABB62169 standard; Protein; 619 AA.		
XX			
AC	ABB62169;		
XX			
DT	26-MAR-2002 (first entry)		
XX			

DE	Drosophila melanogaster polypeptide SEQ ID NO 13299.
XX	
KW	Drosophila: developmental biology; cell signalling; insecticide;
XX	pharmaceutical.
OS	Drosophila melanogaster.
PN	w0200171042-A2.
PD	27-SEP-2001.
XX	
PE	23-MAR-2001: 2001WO-US09231.
XX	
PR	23-MAR-2000: 2000US-191637P.
XX	
PR	11-JUL-2000: 2000US-0614150.
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PMD, Myers EW:
XX	
DR	WPI: 2001-656860/75.
DR	N-PDB; ABL06272.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
PT	
XX	
PS	Disclosure; SEQ ID NO 13299; 21np + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLI6176-ABL30511), expressed DNA sequences (ABLI01840-ABLI6175) and the encoded proteins (ABR57737-ABR7072).
CC	
CC	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
CC	
XX	
SQ	Sequence 619 AA:
Query Match	80.0%; Score 36; DB 22; Length 619;
Best Local Similarity	75.0%; Pred No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0	
QY	1 DGNVPVPM 8 207 dgkvyplm 214
Db	
RESULT 11	
AAY70519	
ID	AAY70519 standard; Protein; 837 AA.
XX	
AC	AAY70519;
XX	
DT	04-JUL-2000 (first entry)
XX	
DE	Clostridium thermocellum xylanase Z.
XX	
KW	Xylanase Z; XynZ; cellulosome; feruloyl esterase; phenolic acid esterase; thermotable; ferulic acid; wheat bran; agricultural byproduct; treat;
KW	grass; paper and pulp industry; feed processing; food additive;
KW	plant cell wall material; degradation.
XX	
OS	Clostridium thermocellum.
XX	
TH	Key Location/Qualifiers
FT	287..296
FT	/label=Pro-/Thr_-region

```

FT Binding-site 289..400
FT /label= Family_VI_cellulose_binding_domain
FT Region 430..453
FT /note= "conserved region"
FT Region 464..487
FT /note= "conserved region"
FT Domain 20..286
FT /note= "This region is a part of multi-domain structure
FT of XynZ. It is specifically claimed and
FT contains feruloyl esterase"
FT Domain 20..307
FT /note= "This region is a part of multi-domain structure
FT of XynZ. It is specifically claimed and
FT contains feruloyl esterase"
FT Domain 20..421
FT /note= "This region is a part of multi-domain structure
FT of XynZ. It is specifically claimed and
FT contains feruloyl esterase"
FT Domain 69..286
FT /note= "specifically exemplified feruloyl esterase"
FT Domain 69..307
FT /note= "specifically exemplified feruloyl esterase"
FT Domain 69..421
FT /note= "specifically exemplified feruloyl esterase"
FT Domain 69..421
FT /note= "specifically exemplified feruloyl esterase"
XX WO200014243-A1.
XX 16-MAR-2000.
XX 03-SEP-1999; 99WO-US20304.
XX 04-SEP-1998; 98US-0099136.
XX (UYGE-) UNIV GEORGIA RES FOUND INC.
XX Blum DL, Kataeva I, Li X, Ljungdahl LG;
XX WPI; 2000-256991/22.
XX N-PSDB; AAZ51818.
XX New recombinant DNA molecule comprising a sequence encoding feruloyl
XX esterase protein, useful for treating grasses and other plant materials
XX used in pulp and paper industries, feed processing and food additives -
XX Claim 23; Pages 37-38; 105pp; English.
XX The present sequence is a xylanase 2 (XynZ) from
XX Clostridium thermocellum. XynZ is an enzymatic component of
XX C. thermocellum cellulosome and has a multi-domain structure which
XX includes a dockerin domain, a catalytic xylanase domain
XX a family VI cellulose binding domain and a domain of
XX unknown function. The unknown domain in the N-terminal region of XynZ
XX has been found to contain feruloyl (phenolic acid) esterase which
XX is involved in the degradation of plant cell wall material. The novel
XX feruloyl esterase is thermostable, easy to purify, has high temperature
XX optima and stable over a wide pH range. The enzyme is used for
XX producing ferulic acid from wheat bran or agricultural byproducts,
XX treating grasses or other plant materials used in the pulp and paper
XX industries, in feed processing and as a food additive.
SQ Sequence 837 AA;

Query Match 80.0%; Score 36; DB 21; Length 837;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 DGNVPVP 7
DB 284 dgnvpvp 290

```

```

AAR30639
ID AAR30639 standard; peptide; 64 AA.
XX AAR30639;
AC AAR30639;
DT 06-MAY-1993 (first entry)
DE Epitope of human CENP-B.
XX Centromere protein; antibodies; polypeptide.
XX Homo sapiens.
XX JP04334398-A.
XX 20-NOV-1992.
XX 08-MAY-1991; 91JP-0102517.
XX 08-MAY-1991; 91JP-0102517.
XX (DAIK ) DAIKIN KOGYO KK.
XX WPI; 1993-005542/01.
XX Human centromere antigen polypeptide - for detection of human
XX antibodies and identification of disease
XX Claim 1; Page 2; 15pp; Japanese.
XX The polypeptide constitutes an epitope of the human centromere protein
XX B (CENP-B) and can be produced by standard recombinant DNA techniques.
XX The polypeptide can be used to detect anti-human centromere antibodies.
XX The type of diseases of a patient having the antibody can be exactly
XX classified using the polypeptide.
XX See also AAR30640-I.
XX Sequence 64 AA;

Query Match 77.8%; Score 35; DB 14; Length 64;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 DGNVPVP 7
DB 1 dgdvpvp 7

RESULT 13
ABG19748
ID ABG19748 standard; Protein; 380 AA.
XX ABG19748;
AC ABG19748;
DT 18-FEB-2002 (first entry)
DE Novel human diagnostic protein #19739.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.

```

RESULT 12

PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB; AAS83935.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 XX Claim 20: SEQ ID No 50107; 103pp: English.
 PS
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 380 AA;
 XX
 Query Match 75.6%; Score 34; DB 22; Length 380;
 Best Local Similarity 75.0%; Pred. No. 2.6e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 DGNVPVPM 8
 II I I I I I
 Db 293 dgmvpmpm 300
 XX
 RESULT 14
 ABB67737
 ID ABB67737 standard; Protein: 585 AA.
 AC ABB67737;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 30003.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KM
 OS Drosophila melanogaster.
 OS
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.

XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL11840.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX Disclosure; SEQ ID NO 30003; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 585 AA;
 XX
 Query Match 75.6%; Score 34; DB 22; Length 585;
 Best Local Similarity 71.4%; Pred. No. 4.1e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 2 GNVVPVPM 8
 II I I I I I
 Db 144 gnapipm 150
 XX
 RESULT 15
 ABG07769
 ID ABG07769 standard; Protein: 759 AA.
 AC ABG07769;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #7760.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 OS
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PA
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB; AAS71956.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 XX Claim 20: SEQ ID No 38128; 103pp: English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. Abg00010-Abg30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 759 AA;

Query Match 75.6%; Score 34; DB 22; Length 759;
Best Local Similarity 71.4%; Pred. No. 5.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 DGNVPVP 7
||:|:|:|
Db 402 dgsyvpip 408

Search completed: June 13, 2002, 08:53:31
Job time: 281 sec

•
•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:54:32 ; Search time 57.95 seconds
(without alignments)
3.372 Million cell updates/sec

Title: US-09-710-339-2_COPY_448_455

Perfect score: 41

Sequence: 1 YTAGQQLT 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/prodata/2/iaa/5A.COMB.pep:*
3: /cgn2_6/prodata/2/iaa/5B.COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6A.COMB.pep:*
5: /cgn2_6/prodata/2/iaa/6B.COMB.pep:*
6: /cgn2_6/prodata/2/iaa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	478	1	US-08-720-899-7
2	41	100.0	478	1	US-08-459-610-7
3	41	100.0	478	2	US-08-343-804-7
4	41	100.0	478	2	US-08-339-715A-2
5	41	100.0	478	2	US-08-600-908A-10
6	41	100.0	478	3	US-08-683-838A-10
7	41	100.0	478	4	US-09-182-859-7
8	31	75.6	342	1	US-08-499-215-2
9	30	73.2	109	2	US-08-585-585A-1
10	30	73.2	109	2	US-08-249-037C-1
11	30	73.2	109	2	US-08-788-622B-1
12	30	73.2	109	3	US-08-788-622B-1
13	30	73.2	631	3	US-08-814-052-8
14	30	73.2	631	3	US-08-812-829-8
15	29	70.7	177	3	US-08-824-692-30
16	29	70.7	229	3	US-08-824-692-31
17	29	70.7	290	3	US-08-824-692-29
18	29	70.7	505	1	US-08-220-603A-10
19	29	70.7	1463	1	US-08-220-603A-11
20	28	68.3	441	1	US-08-403-866-10
21	27	65.9	214	4	US-08-960-507-21
22	27	65.9	215	4	US-09-247-373B-40
23	27	65.9	263	1	US-08-407-544-2
24	27	65.9	264	1	US-08-719-697-8
25	27	65.9	264	4	US-08-727-616A-8
26	27	65.9	265	2	US-08-719-697-2
27	27	65.9	265	2	US-08-719-697-2
27	27	65.9	265	2	US-08-719-697-2

ALIGNMENTS

28	27	65.9	265	4	US-08-727-616A-2	Sequence 2, Appli
29	27	65.9	265	4	US-08-727-616A-6	Sequence 6, Appli
30	27	65.9	286	1	US-07-721-775A-2	Sequence 2, Appli
31	27	65.9	286	1	US-08-339-658-2	Sequence 2, Appli
32	27	65.9	286	1	US-08-346-333-16	Sequence 16, Appli
33	27	65.9	286	2	US-08-719-697-4	Sequence 4, Appli
34	27	65.9	286	4	US-09-263-933-7	Sequence 7, Appli
35	27	65.9	286	4	US-09-263-933-14	Sequence 14, Appli
36	27	65.9	286	4	US-08-727-616A-4	Sequence 21, Appli
37	27	65.9	286	4	US-09-025-769B-285	Sequence 265, App
38	27	65.9	286	4	US-09-025-769B-362	Sequence 362, App
39	27	65.9	286	5	PCT-US91-0750B-16	Sequence 16, Appli
40	27	65.9	299	4	US-09-025-769B-285	Sequence 285, App
41	27	65.9	299	4	US-09-025-769B-298	Sequence 298, App
42	27	65.9	299	4	US-09-025-769B-300	Sequence 300, App
43	27	65.9	406	4	US-09-108-020-4	Sequence 4, Appli
44	27	65.9	406	4	US-09-108-020-4	Sequence 4, Appli
45	27	65.9	406	4	US-09-108-020-4	Sequence 42, Appli

RESULT 1
US-08-720-899-7
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720, 899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343, 804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowrey Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054, 214-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-720-899-7

Query Match 100.0%; Score 41; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
|||||
Db 428 YTAGOQLT 435

RESULT 2
US-08-459-610-7
; Sequence 7, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NO. 58010430 No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459, 610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-459-610-7

Query Match 100.0%; Score 41; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
|||||
Db 428 YTAGOQLT 435

RESULT 3
US-08-343-804-7
; Sequence 7, Application US/08343804
; Patent No. 5830837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borcherdt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS

NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NO. 58308370 No. 5830837disk of No. 5830837th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,804
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-343-804-7

Query Match 100.0%; Score 41; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
|||||
Db 428 YTAGOQLT 435

RESULT 4
US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michiyo
; APPLICANT: Okada, Shigetaka
; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POPHAM, HAIK, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; STREET: G. Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,715A
; FILING DATE: 14-NOV-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Meiser, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335.009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 100.0%; Score 41; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOLT 8
DB 427 YTAGOOLT 434

RESULT 5
US-08-600-908A-10
Sequence 10, Application US/08600908A
Patent No. 5989169
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59891690 No. 5989169disk of No. 5989169th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,908A
FILING DATE: 13-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ. ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-600-908A-10

Query Match 100.0%; Score 41; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOLT 8
DB 428 YTAGOOLT 435

RESULT 6
US-08-683-838A-10
Sequence 10, Application US/08683838A
Patent No. 6022724
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60227240 No. 6022724disk of No. 6022724th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,838A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/600,908
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ. ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-683-838A-10

Query Match 100.0%; Score 41; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOLT 8
DB 428 YTAGOOLT 435

RESULT 7
US-09-182-859-7
Sequence 7, Application US/09182859
Patent No. 6143708
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Borchert, Torben
APPLICANT: Bisgard-Frantzen, Henrik

;; TITLE OF INVENTION: Alpha-Amylase Mutants
;; FILE REFERENCE: 4796.204-US
;; CURRENT APPLICATION NUMBER: US/09/182,859
;; CURRENT FILING DATE: 1998-10-29
;; EARLIER APPLICATION NUMBER: 0515/96
;; EARLIER FILING DATE: 1996-04-30
;; EARLIER APPLICATION NUMBER: 0712/96
;; EARLIER FILING DATE: 1996-06-28
;; EARLIER APPLICATION NUMBER: 0775/96
;; EARLIER FILING DATE: 1996-07-11
;; EARLIER APPLICATION NUMBER: 1263/96
;; EARLIER FILING DATE: 1996-11-08
;; NUMBER OF SEQ ID NOS: 37
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 7
;; LENGTH: 478
;; TYPE: PRI
;; ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match 100.0%; Score 41; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGQOLT 8
Db 428 YTAGQOLT 435

RESULT 8
US-08-499-215-2
; Sequence 2, Application US/08499215
; Patent No. 5612204
; GENERAL INFORMATION:
; APPLICANT: Saeki, Hisashi
; APPLICANT: Miura, Akira
; TITLE OF INVENTION: BIOLOGICAL DEGRADATIVE
; TITLE OF INVENTION: TREATMENT OF CHLORINE-SUBSTITUTED ETHYLENE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Miron, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue
; CITY: N.W.
; STATE: Washington D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/499,215
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP Hei-6-179689
; FILING DATE: 08-JUL-1994
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 342 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-499-215-2

Query Match 75.6%; Score 31; DB 1; Length 342;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YTAGQOLT 8
Db 34 YTAGQOLT 41

RESULT 9
US-08-585-585A-1
; Sequence 1, Application US/08585585A
; Patent No. 5874308
; GENERAL INFORMATION:
; APPLICANT: Kilburn, Douglas G.
; APPLICANT: Humphries, Keith R.
; APPLICANT: Doherty, James G.
; APPLICANT: Jervis, Eric
; APPLICANT: Alimonti, Judie
; TITLE OF INVENTION: Compositions and methods for modulating
; TITLE OF INVENTION: cell proliferation using growth factor-polysaccharide
; TITLE OF INVENTION: binding fusion proteins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rae-Venter Law Group
; STREET: Box 60039
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/585,585A
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: King, Viola T.
; REGISTRATION NUMBER: 41,131
; REFERENCE/DOCKET NUMBER: CBPT.016.0005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)328-4400
; TELEFAX: (650)328-4477
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-585-585A-1

Query Match 73.2%; Score 30; DB 2; Length 109;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGQOLT 7
Db 40 YTAGQOLT 46

RESULT 10
US-08-249-037C-1
; Sequence 1, Application US/08249037C
; Patent No. 5928917
; GENERAL INFORMATION:
; APPLICANT: Kilburn, Douglas G.
; APPLICANT: Miller, Robert C.
; APPLICANT: Warren, Richard A.J.
; APPLICANT: Gilkes, Neil R.
; TITLE OF INVENTION: Polysaccharide binding fusion proteins
; TITLE OF INVENTION: and conjugates
; NUMBER OF SEQUENCES: 21

```

CORRESPONDENCE ADDRESS:
ADDRESS: Rae-Venter Law Group, P.C.
STREET: P.O. Box 60039
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,037C
FILING DATE: 24-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/865,095
FILING DATE: 08-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,987
FILING DATE: 25-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/216,794
FILING DATE: 08-JUL-1988
ATTORNEY/AGENT INFORMATION:
NAME: Kung, Viola T.
REGISTRATION NUMBER: 41,131
REFERENCE/DOCKET NUMBER: CBDT.002.04US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)328-4400
TELEFAX: (650)328-4477
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-249-037C-1

Query Match      73.2%; Score 30; DB 2; Length 109;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOL 7
DB 40 YTAGORI 46

RESULT 11
US-08-788-622B-1
Sequence 1, Application US/08788622B
Patent No. 5962289
GENERAL INFORMATION:
APPLICANT: Kilburn, Douglas G.
APPLICANT: Miller, Robert C.
APPLICANT: Warren, Richard A.J.
APPLICANT: Gilkes, Neil R.
TITLE OF INVENTION: Polysaccharide binding fusion proteins
TITLE OF INVENTION: and conjugates
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rae-Venter Law Group, P.C.
STREET: P.O. Box 60039
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/788,622B
FILING DATE: January 23, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/249,037
FILING DATE: 24-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/865,095
FILING DATE: 08-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,987
FILING DATE: 25-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/216,794
FILING DATE: 08-JUL-1988
ATTORNEY/AGENT INFORMATION:
NAME: Kung, Viola T.
REGISTRATION NUMBER: 41,131
REFERENCE/DOCKET NUMBER: CBDT.002.06US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)328-4400
TELEFAX: (650)328-4477
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-788-622B-1

Query Match      73.2%; Score 30; DB 2; Length 109;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOL 7
DB 40 YTAGORI 46

RESULT 12
US-08-788-621B-1
Sequence 1, Application US/08788621B
Patent No. 6124117
GENERAL INFORMATION:
APPLICANT: Kilburn, Douglas G.
APPLICANT: Miller, Robert C.
APPLICANT: Warren, Richard A.J.
APPLICANT: Gilkes, Neil R.
TITLE OF INVENTION: Polysaccharide binding fusion proteins
TITLE OF INVENTION: and conjugates
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rae-Venter Law Group, P.C.
STREET: P.O. Box 60039
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/788,621B
FILING DATE: January 23, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/249,037
FILING DATE: 24-MAY-1994

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/865,095
FILING DATE: 08-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,987
FILING DATE: 25-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/216,794
FILING DATE: 08-JUL-1988
ATTORNEY/AGENT INFORMATION:
NAME: Kung, Viola T.
REGISTRATION NUMBER: 41,131
REFERENCE/DOCKET NUMBER: CDDT.002.0505
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)328-4400
TELEFAX: (650)328-4477
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-788-621B-1

Query Match 73.2%; Score 30; DB 3; Length 109;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOOL 7
|||||:
DB 40 YTAGORI 46

RESULT 13
US-08-814-052-8
Sequence 8, Application US/08814052
Patent No. 6015783
GENERAL INFORMATION:
APPLICANT: von der Osten, Claus
APPLICANT: Cherry, Joel R.
APPLICANT: Bjornvad, Mads E.
APPLICANT: Vind, Jesper
APPLICANT: Rasmussen, Michael Dolberg
TITLE OF INVENTION: PROCESS FOR REMOVAL OR BLEACHING OF SOILING
TITLE OF INVENTION: OR STRAINS FROM CELLULOSIC FABRIC
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSER: No. 60157830 No. 6015783disk of No. 6015783th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/814,052
FILING DATE: 06-MAR-1997
CLASSIFICATION: 510
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4684.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 631 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 6015783e
US-08-814-052-8

Query Match 73.2%; Score 30; DB 3; Length 631;
Best Local Similarity 71.4%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOOL 7
|||||:
DB 73 YTAGORI 79

RESULT 14
US-08-812-829-8
Sequence 8, Application US/08812829
Patent No. 6017751
GENERAL INFORMATION:
APPLICANT: von der Osten, Claus
APPLICANT: Bjornvad, Mads E.
APPLICANT: Vind, Jesper
APPLICANT: Rasmussen, Michael Dolberg
TITLE OF INVENTION: PROCESS FOR REMOVAL OR BLEACHING OF SOILING
TITLE OF INVENTION: OR STRAINS FROM CELLULOSIC FABRIC
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSER: No. 60177510 No. 6017751disk of No. 6017751th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812,829
FILING DATE: 06-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4690.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 631 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6017751e
US-08-812-829-8

Query Match 73.2%; Score 30; DB 3; Length 631;
Best Local Similarity 71.4%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOOL 7
|||||:
DB 73 YTAGORI 79

```

RESULT 15
US-08-824-692-30
; Sequence 30, Application US/08824692
; Patent No. 6017703
; GENERAL INFORMATION:
; APPLICANT: Kinders, Robert J.
; APPLICANT: Enfield, David L.
; APPLICANT: Hass, G. Michael
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING
; TITLE OF INVENTION: FOR OR MODULATING A TUMOR ASSOCIATED ANTIGEN
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,692
; FILING DATE: 08-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 130001.404
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 622-6031
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-824-692-30

```

```

Query Match 70.7%; Score 29; DB 3; Length 177;
Best Local Similarity 62.5%; Pred. No. 37;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 YTAGOOLT 8
| | | | |
DB 63 YKAGEOVT 70

```

Search completed: June 13, 2002, 08:54:33
Job time: 318 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:15 ; Search time 125.81 Seconds

(Without alignments)
11.000 Million cell updates/sec

Title: US-09-710-339-2_COPY_448_455

Perfect score: 41

Sequence: 1 YTAGQQLT 8

Scoring table: BLOSUM62

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	499	3 Q96TH4	Q96TH4 aspergillus
2	34	82.9	849	16 Q9KND8	Q9KND8 vibrio chol
3	33	80.5	420	16 Q982J7	Q982J7 rhizobium 1
4	32	78.0	192	12 Q9YWK5	Q9YWK5 buzura supp
5	32	78.0	211	5 Q9GPI6	Q9GPI6 anopheles g
6	32	78.0	333	12 Q99H12	Q99H12 heliocoverpa
7	32	78.0	390	12 Q9YWS6	Q9YWS6 lymantria d
8	32	78.0	394	12 Q9PYW5	Q9PYW5 xestia c-ni
9	32	78.0	399	12 Q9JML8	Q9JML8 choristoneu
10	32	78.0	399	12 Q92407	Q92407 bombyx mori
11	32	78.0	400	12 Q9JGL4	Q9JGL4 epiphyas po
12	32	78.0	400	12 Q9J822	Q9J822 spodoptera
13	32	78.0	412	12 Q9J8X7	Q9J8X7 heliocoverpa
14	32	78.0	422	12 Q9J8I5	Q9J8I5 spodoptera
15	32	78.0	460	12 Q9JEX7	Q9JEX7 cydia pomon
16	32	78.0	541	5 Q93W09	Q93W09 glardia lam

17	32	78.0	764	4 Q9P2H3	Q9P2H3 homo sapien
18	32	78.0	1250	5 Q23409	Q23409 caenorhabdi
19	32	78.0	2119	13 Q90X47	Q90X47 brachydanio
20	31	75.6	214	2 Q9LOE0	Q9LOE0 streptomyce
21	31	75.6	221	16 Q9JUN7	Q9JUN7 neisseria m
22	31	75.6	223	5 Q9V8H2	Q9V8H2 drosophila
23	31	75.6	242	9 Q03913	Q03913 bacteriopho
24	31	75.6	271	16 Q98BX7	Q98BX7 rhizobium 1
25	31	75.6	271	16 Q92NE0	Q92NE0 rhizobium m
26	31	75.6	302	5 Q9W335	Q9W335 drosophila
27	31	75.6	343	2 Q53025	Q53025 rhodococcus
28	31	75.6	347	16 Q9KGN0	Q9KGN0 bacillus ha
29	31	75.6	386	12 Q9DXY2	Q9DXY2 pluteella xy
30	30	73.2	52	9 Q9B036	Q9B036 mycobacteri
31	30	73.2	99	11 Q9ERU1	Q9ERU1 talus norv
32	30	73.2	104	4 Q9NTE2	Q9NTE2 homo sapien
33	30	73.2	108	17 Q9YFY0	Q9YFY0 aeropyrum p
34	30	73.2	146	13 Q9DF33	Q9DF33 ophiophagus
35	30	73.2	222	5 Q9V8H1	Q9V8H1 drosophila
36	30	73.2	222	5 Q9V8H0	Q9V8H0 drosophila
37	30	73.2	238	5 Q9Y098	Q9Y098 periplaneta
38	30	73.2	240	5 Q9V8G5	Q9V8G5 drosophila
39	30	73.2	283	16 Q9AAJ7	Q9AAJ7 caulobacter
40	30	73.2	331	16 Q9HYE6	Q9HYE6 pseudomonas
41	30	73.2	388	16 Q9RSN6	Q9RSN6 deinococcus
42	30	73.2	410	2 Q9XEV7	Q9XEV7 streptomyce
43	30	73.2	425	17 Q59010	Q59010 pyrococcus
44	30	73.2	452	10 Q9ZQ00	Q9ZQ00 arabidopsis
45	30	73.2	466	6 Q95KH6	Q95KH6 macaca fasc

ALIGNMENTS

RESULT 1
ID Q96TH4 PRELIMINARY: PRT: 499 AA.
AC Q96TH4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TAKA-AMYLASE A (EC 3.2.1.1) (ALPHA-AMYLASE).
GN AMYA OR AMY1.
OS Aspergillus oryzae, and
OS Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062, 5059;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=A.oryzae; STRAIN=R1840;
RX MEDLINE=20289310; PubMed=10830498;
RA Gomi K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N.,
RA Ilanura Y.;
RT "Molecular cloning and characterization of a transcriptional activator
gene, amyR, involved in the amylolytic gene expression in Aspergillus
oryzae";
RT Biosci. Biotechnol. Biochem. 64:816-827(2000).
RL [2]
RN SEQUENCE FROM N.A.
RP SPECIES=A.flavus; STRAIN=86-10D;
RC Fakhoury A.M., Moloshuk C.P.;
RA "Amyl, the alpha-amylase gene of Aspergillus flavus: Involvement in
afatoxin biosynthesis in maize kernels";
RT Phytopathology 89:908-914(1999).
RL EMBL: AB021876; BA95703.1; -;
DR EMBL: AF139925; AF14264.1; -;
SQ SEQUENCE 499 AA; 54804 MW; EEFA2ADA71D20DA9 CRC64;

Query Match 100.0%; Score 41; DB 3; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOOLT 8
 |||||
 Db 449 YTAGOOLT 456

RESULT 2

O9KND8 PRELIMINARY; PRT; 849 AA.
 AC O9KND8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CHITINASE.
 GN VCA0027.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 NCBI_TaxID=66;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., McKelanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae.";
 RT Nature 406:477-483(2000).
 CC -1. SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 HYDROLASES).
 DR EMBL: AE004346; AAF95941.1; -.
 DR HSSP: P07254; ICTN.
 DR TIGR: VCA0027; -.
 DR InterPro: IPR001579; Chitinase_2.
 DR InterPro: IPR003610; Chitin_bind3.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR InterPro: IPR000601; PKD domain.
 DR Pfam: PF00704; Glyco_hydro_18; 1.
 DR Pfam: PF00801; PKD; 2.
 DR SMART: SM00495; ChtdB3; 1.
 DR SMART: SM00089; PKD; 2.
 DR PROSITE: PS01095; CHITINASE_18; 1.
 DR Complete proteome: Glycosidase; Hydrolase.
 KW SEQUENCE 849 AA; 90314 MW; DB164D55673FD34 CRC64;

Query Match 82.9%; Score 34; DB 16; Length 849;
 Best Local Similarity 75.0%; Pred. No. 57;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGOOLT 8
 |||||
 Db 805 YTAGOOLT 812

RESULT 3

O982J7 PRELIMINARY; PRT; 420 AA.
 AC O982J7;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE ABC TRANSPORTER, PERIPLASMIC PROTEIN.
 GN MUR9034.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Plasmid pMA.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 NCBI_TaxID=381;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;

RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimizu S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP003015; BAB54459.1; -.
 KW Plasmid; Complete proteome.
 SQ SEQUENCE 420 AA; 45026 MW; 95E3EF93E2CB093 CRC64;

Query Match 80.5%; Score 33; DB 16; Length 420;
 Best Local Similarity 75.0%; Pred. No. 45;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGOOLT 8
 |||||
 Db 83 YTAGOOLT 90

RESULT 4

O9YWK5 PRELIMINARY; PRT; 192 AA.
 AC O9YWK5;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE HYPOTHETICAL 22.7 KDA PROTEIN (FRAGMENT).
 OS Buzura suppressaria nuclear polyhedrosis virus (BSNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=74320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99036048; PubMed=9820162;
 RA Hu Z.H., Arif B.M., Jin F., Martens J.W.M., Chen X.W., Sun J.S.,
 RA Zuidema D., Goldbach R.W., Vlak J.M.;
 RT "Distinct gene arrangement in the Buzura suppressaria single-
 nucleocapsid nucleopolyhedrovirus genome.";
 RL J. Gen. Virol. 79:2841-2851(1998).
 DR EMBL: AF058929; AAC77811.1; -.
 KW Hypothetical protein.
 FT NON_TER 192
 SQ SEQUENCE 192 AA; 22713 MW; D974F7B02895DC8E CRC64;

Query Match 78.0%; Score 32; DB 12; Length 192;
 Best Local Similarity 75.0%; Pred. No. 32;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGOOLT 8
 |||||
 Db 141 YTAGOOLT 148

RESULT 5

O9GPL6 PRELIMINARY; PRT; 211 AA.
 AC O9GPL6;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE GLUTATHIONE S-TRANSFERASE 1-10.
 OS Anopheles gambiae (African malaria mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 OC Anophelinae.

OX NCBI_TaxID=7165;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ZAN/U.
 RA Orrelli F., Ranson H., Jensen B., Rossiter L., Wang X., Roth C.,
 RA Collins F.H., Hemingway J.;
 RT "Identification of a novel class of insect glutathione S-transferases
 RT associated with DDT resistance in the malaria vector *Anopheles*
 RT *gambiae*.";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF316638; AAC5166.1; -
 DR InterPro: IPR004046; GST_C.
 DR InterPro: IPR004045; GST_N.
 KW Transferase.
 SQ SEQUENCE 211 AA; 24439 MW; F37363CE18ADE71 CRC64;

Query Match 78.0%; Score 32; DB 5; Length 211;
 Best Local Similarity 75.0%; Pred. No. 36;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
 Db 144 YVAGQKLT 151

RESULT 6
 ID 099H12 PRELIMINARY; PRT; 333 AA.
 AC 099H12;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE P47.
 OS *Helicoverpa armigera* nucleopolyhedrovirus G4.
 CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC Nucleopolyhedrovirus.
 OX NCBI_TaxID=148363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Deng F., Chen X., Vlask J.M., Arif B.M., Hu Z.;
 RT "Sequence analysis of the gp37 gene of *Heliothis armigera* single-
 RT nucleocapsid nucleopolyhedrovirus.";
 RL Zhongguo Bingduxue 15:35-42(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Wang H., Hu Z., Sun X., Vlask J.M., Chen X.;
 RT "Sequence analysis of the iap3 gene of *Heliothis armigera* single-
 RT nucleocapsid nucleopolyhedrovirus.";
 RL Zhongguo Bingduxue 15:43-49(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21078302; PubMed=11210934;
 RA Wang H., Chen X., Wang H., Arif B.M., Vlask J.M., Hu Z.;
 RT "Nucleotide sequence and transcriptional analysis of a putative basic
 RT DNA-binding protein of *Helicoverpa armigera* polyhedrovirus.";
 RL Virus Genes 22:113-120(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21064569; PubMed=11125177;
 RA Chen X., Ijkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,
 RA Peters S., Zuidema D., Lankhorst R.K., Vlask J.M., Hu Z.;
 RT "The sequence of the *Helicoverpa armigera* single-nucleocapsid
 RT nucleopolyhedrovirus genome.";
 RL J. Gen. Virol. 82:241-257(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Chen X., Ijkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,
 RA Peters S., Zuidema D., Lankhorst R.K., Vlask J.M., Hu Z.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF211059; AAC53778.1; -
 SQ SEQUENCE 333 AA; 38963 MW; B8EBDAF391A40285 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 333;
 Best Local Similarity 75.0%; Pred. No. 59;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
 Db 161 YTLGOQLS 168

RESULT 7
 ID 09YMS6 PRELIMINARY; PRT; 390 AA.
 AC 09YMS6;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE LP-P47.
 OS *Lymantria dispar* multicapsid nuclear polyhedrosis virus (LdMNPV).
 CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10449;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99124765; PubMed=9867315;
 RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
 RA Slavicek J.M., Rohmann G.F.;
 RT "Sequence and analysis of the genome of a baculovirus pathogenic for
 RT *Lymantria dispar*.";
 RL Virology 253:17-34(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kuzio J.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF081610; AAC70233.1; -
 SQ SEQUENCE 390 AA; 45535 MW; 1A982632B037C552 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 390;
 Best Local Similarity 75.0%; Pred. No. 69;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
 Db 140 YTLGOQLS 147

RESULT 8
 ID 09PYW5 PRELIMINARY; PRT; 394 AA.
 AC 09PYW5;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ORF78.
 OS *Xestia c-nigrum* granulosis virus (XngV) (*Xestia c-nigrum*
 OS granulovirus).
 CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
 OX NCBI_TaxID=51677;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99434230; PubMed=10502508;
 RA Hayakawa T., Ko R., Okano K., Seong S.I., Goto C., Maeda S.;
 RT "Sequence analysis of the *Xestia c-nigrum* granulovirus genome.";
 RL Virology 262:277-297(1999).
 DR EMBL: AF162221; AAF05192.1; -
 SQ SEQUENCE 394 AA; 46157 MW; 4C598BFCB5235DBF CRC64;

Query Match 78.0%; Score 32; DB 12; Length 394;
 Best Local Similarity 75.0%; Pred. No. 70;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
 111111:
 Db 141 YTLGOQLS 148

RESULT 9
 09IML8 PRELIMINARY; PRT; 399 AA.

AC 09IML8; MEDLINE=99281911; PubMed=10355780;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE P47.
 GN P47.
 OS Choristoneura fumiferana nuclear polyhedrosis virus (CFMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10448;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EC-1;
 RX MEDLINE=20276145; PubMed=10814576;
 RA Lapointe R., Back D.W., Ding Q., Carstens E.B.;
 RT "Identification and molecular characterization of the choristoneura
 fumiferana multicausid nucleopolyhedrovirus genomic region encoding
 the regulatory genes pkip, p47, lef-12, and gta.";
 RL Virology 271:109-121(2000).
 DR EMBL: AF177329; AAF64092.1; -
 SQ SEQUENCE 399 AA; 46323 MW; 744600382F7DC942 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 399;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
 111111:
 Db 148 YTLGOQLS 155

RESULT 10
 092407 PRELIMINARY; PRT; 399 AA.

AC 092407; MEDLINE=97329351; PubMed=9185864;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE P47-ACMNPV ORF40.
 GN P47.
 OS Bombyx mori nuclear polyhedrosis virus (BmNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10458;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T3;
 RX MEDLINE=97329351; PubMed=9185864;
 RA Kamita S.G., Maeda S.;
 RT "Sequencing of the putative DNA helicase-encoding gene of the Bombyx
 mori nuclear polyhedrosis virus and fine-mapping of a region involved
 in host range expansion.";
 RL Gene 190:173-179(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T3;
 RX MEDLINE=99281911; PubMed=10355780;
 RA Gomi S., Majima K., Maeda S.;
 RT "Sequence analysis of the genome of Bombyx mori
 nucleopolyhedrovirus.";
 RL J. Gen. Virol. 80:1323-1337(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T3;

RA Maeda S.;
 RL Submitted (May-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T3;
 RA Gomi S.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U3180; AAC63715.1; -
 SQ SEQUENCE 399 AA; 47260 MW; 69114411B659C63 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 399;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
 111111:
 Db 144 YTLGOQLS 151

RESULT 11
 091G14 PRELIMINARY; PRT; 399 AA.

AC 091G14; MEDLINE=20036646; PubMed=10567663;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE P47.
 GN P47.
 OS Epiphyas postvittana nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=70600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hyink O., Dellow R.A., Olsen M., Caradoc-Davies K.M.B., Drake K.,
 RA Ward V.K.;
 RT "The complete sequence of the Epiphyas postvittana
 nucleopolyhedrovirus genome.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY043265; AAK85601.1; -
 SQ SEQUENCE 399 AA; 46410 MW; 6495E5250ABFAF11 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 399;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
 111111:
 Db 148 YTLGOQLS 155

RESULT 12
 09J822 PRELIMINARY; PRT; 400 AA.

AC 09J822; MEDLINE=20036646; PubMed=10567663;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE ORF115 P47.
 OS Spodoptera exigua nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10454;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20036646; PubMed=10567663;
 RA IJdel W.F., van Strien E.A., Heidens J.G., Broer R., Zuidema D.,
 RA Goldbach R.W., Vlak J.M.;
 RT "Sequence and organization of the spodoptera exigua multicausid
 nucleopolyhedrovirus genome.";
 RL J. Gen. Virol. 80:3289-3304(1999).

RN [2]
 RT SEQUENCE FROM N.A.
 RA Jukel W.F.J., van Strien E.A., Heldens J.G.M., Broer R., Zuidema D.,
 RA Goldbach R.W., Vlak J.M.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF169823; AAF33644.1; -
 SO SEQUENCE 400 AA; 46896 MW; 0373F30953553F5B CRC64;

Query Match 78.0%; Score 32; DB 12; Length 400;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 YTAGOOLT 8
 Db 144 YTLGQOLS 151

RESULT 13
 Q91BX7 PRELIMINARY; PRT; 412 AA.
 AC Q91BX7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE P47.
 OS Helicoverpa armigera nuclear polyhedrosis virus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=51313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C1;
 RA Zhang C.X., Wu J.C.;
 RT "Genome structure and the p10 gene of the Helicoverpa armigera
 RT nucleopolyhedrovirus.";
 RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 33:179-184(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C1;
 RA Zhang C.X., Jin W.R.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF303045; AAK96288.1; -
 SO SEQUENCE 412 AA; 48091 MW; DBA6452E9159AA85 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 412;
 Best Local Similarity 75.0%; Pred. No. 74;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 YTAGOOLT 8
 Db 161 YTLGQOLS 168

RESULT 14
 Q91B15 PRELIMINARY; PRT; 422 AA.
 AC Q91B15;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE TRANSCRIPTION REGULATOR P47.
 OS Spodoptera litura nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=46242;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=G2;
 RX MDLINE=21425398; PubMed=11531416;
 RA Pang Y., Yu J., Wang L., Hu X., Bao W., Li G., Chen C., Han H., Hu S.,
 RA Yang H.;

RT *Sequence Analysis of the Spodoptera litura Multicapsid
 RT Nucleopolyhedrovirus Genome.*;
 RL Virology 287:391-404(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=G2;
 RC STRAIN=G2;
 RA Yu J., Wang L., Hu X., Pang Y.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF325155; AAL01722.1; -
 SO SEQUENCE 422 AA; 49358 MW; 64D30486B1522621 CRC64;

Query Match 78.0%; Score 32; DB 12; Length 422;
 Best Local Similarity 75.0%; Pred. No. 75;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 YTAGOOLT 8
 Db 152 YTLGQOLS 159

RESULT 15
 Q91EY7 PRELIMINARY; PRT; 460 AA.
 AC Q91EY7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ORF68 P47.
 GN ORF68 P47.
 OS Cydia pomonella granulosis virus (CPGV) (Cydia pomonella
 OS granulovirus).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
 OX NCBI_TaxID=28289;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RX MEDLINE=93188168; PubMed=8445726;
 RA Crook N.E., Clem R.J., Miller L.K.;
 RT "An apoptosis-inhibiting baculovirus gene with a zinc finger-like
 RT motif.";
 RL J. Virol. 67:2168-2174(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RX MEDLINE=96207404; PubMed=8615018;
 RA Theilmann D.A., Chantler J.K., Stewart S., Flipsen H.T., Vlak J.M.,
 RA Crook N.E.;
 RT "Characterization of a highly conserved baculovirus structural protein
 RT that is specific for occlusion-derived virions.";
 RL Virology 218:148-158(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RX MEDLINE=97380577; PubMed=9237352;
 RA Kang W., Crook N.E., Winstanley D., O'Reilly D.R.;
 RT "Complete sequence and transposon mutagenesis of the BamHI J fragment
 RT of Cydia pomonella granulosis virus.";
 RL Virus Genes 14:131-136(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RX MEDLINE=98418511; PubMed=9747739;
 RA Kang W., Tristem M., Maeda S., Crook N.E., O'Reilly D.R.;
 RT "Identification and characterization of the Cydia pomonella
 RT granulovirus cathepsin and chitinase genes.";
 RL J. Gen. Virol. 79:2283-2292(1998).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RA Luque T., Finch R., Crook N., O'Reilly D.R., Winstanley D.;
 RT "The complete sequence of the Cydia pomonella granulovirus genome.";
 RT Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: U53466; AAK70728.1; -
SO SEQUENCE 460 AA; 54126 MW; 8219FA11B7DD1A6B CRC64;

Query Match 78.0%; Score 32; DB 12; Length 460;
Best Local Similarity 75.0%; Pred. No. 83;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 YTAGGQQLT 8
|||
Db 204 YTLGQQLS 211

Search completed: June 13, 2002, 09:11:17
Job time: 1062 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:54 ; Search time 38.1 Seconds

(without alignments)
8.130 Million cell updates/sec

Title: US-09-710-339-2_COPY_448_455

Perfect score: 41

Sequence: 1 YTAGQQLT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	41	100.0	498	AMYA_ASPAW	Q02905 aspergillus
2	41	100.0	499	AMYA_ASPOR	P10529 aspergillus
3	41	100.0	499	AMYB_ASPAW	Q02906 aspergillus
4	41	100.0	499	AMT_ASPSH	P30292 aspergillus
5	33	80.5	370	NRF7_HAELN	P44942 haemophilus
6	32	78.0	399	VP47_NPVOP	O10300 oryza psen
7	32	78.0	401	VP47_NPVAC	P34051 autographa
8	31	75.6	215	GP21_DROME	Q9vg98 drosophila
9	31	75.6	215	GP23_DROME	Q9vg96 drosophila
10	31	75.6	226	PURQ_STRCO	Q9kkk6 streptomyce
11	31	75.6	282	YV69_VITBO	Q9kkyl vibrio chol
12	31	75.6	442	NX1B_HUMAN	P58400 homo sapien
13	31	75.6	554	CX1A_PARDE	P08305 paracoccus
14	31	75.6	558	CX1B_PARDE	P08002 paracoccus
15	31	75.6	699	CH11_BACCI	P20533 bacillus ci
16	31	75.6	1477	NX1A_HUMAN	Q9ulb1 homo sapien
17	31	75.6	5179	MOC2_HUMAN	Q02817 homo sapien
18	30	73.2	390	DACD_SALTY	P37604 salmonella
19	30	73.2	449	GUNA_CELFI	P07984 cellulomona
20	30	73.2	473	NORR_PSEST	P08008 pseudomonas
21	29	70.7	105	CU12_HYACE	P45568 hyalophora
22	29	70.7	270	CH12_LYCES	P10708 lycopersico
23	29	70.7	270	CH12_LYCES	P10708 lycopersico
24	29	70.7	366	TRPD_MYCLE	P13869 petunia hyb
25	29	70.7	388	DACD_ECOLI	P33013 escherichia
26	29	70.7	466	NORR_PSEAE	O59647 pseudomonas
27	29	70.7	512	YMO9_MYCTU	O10398 mycobacteri
28	29	70.7	551	LIDP_ECOLI	P33231 escherichia
29	29	70.7	565	COX1_RHOSH	P33517 rhodobacter
30	29	70.7	728	EP2_HALHA	P14823 halobacteri
31	29	70.7	735	ABC8_HUMAN	Q9nut2 homo sapien
32	29	70.7	922	CLPA_PEA	P5100 plism sativ
33	29	70.7	923	CLAB_LYCES	P31542 lycopersico

34	29	70.7	926	1	CLAB_LYCES	P31541 lycopersico
35	29	70.7	1024	1	CARC_HUMAN	Q9npp4 homo sapien
36	29	70.7	1231	1	CPAH_HUMAN	P08603 homo sapien
37	29	70.7	1328	1	YMD9_YEAST	P03434 saccharomyc
38	29	70.7	1328	1	YMT5_YEAST	O04214 saccharomyc
39	29	70.7	1375	1	NID2_HUMAN	O14112 homo sapien
40	29	70.7	1463	1	PA2R_BOVIN	P49259 bos taurus
41	28	68.3	87	1	V476_MYCLE	Q9cda3 mycobacteri
42	28	68.3	129	1	RL3_TERPS	P11252 yersinia ps
43	28	68.3	133	1	REV_CAEYC	P33460 caprine art
44	28	68.3	210	1	GTT3_MUSDO	P46432 musca domes
45	28	68.3	217	1	RL3_MYCBO	O06044 mycobacteri

ALIGNMENTS

RESULT	1	AMYA_ASPAW	STANDARD	PRT	498 AA.
ID	AMYA_ASPAW				
AC	Q02905				
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	01-OCT-1996	(Rel. 34, Last annotation update)			
DE	Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan				
DE	glucanohydrolase A).				
OS	AMYA.				
CC	Aspergillus awamori.				
CC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
CC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				
OX	NCBI_Taxid=105351;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=UK143F;				
RX	MEDLINE=90254827; Pubmed=2340591;				
RA	Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,				
RA	Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,				
RA	Berka R.M.;				
RT	"Cloning, characterization, and expression of two alpha-amylase genes				
RT	from Aspergillus niger var. awamori.";				
RL	Curr. Genet. 17:203-212(1990).				
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic				
CC	linkages in oligosaccharides and polysaccharides.				
CC	-1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO				
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-slb.ch/announce/				
CC	or send an email to license@isb-slb.ch).				
DR	EMBL; X52755; CAA36966.1; -				
DR	HSSP; P10529; TTA.				
DR	InterPro; IPR000461; Alpha-amylase.				
DR	Pfam; PF00128; alpha-amylase; 1.				
KW	Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;				
KW	Glycoprotein; Signal; Multigene family.				
FT	SIGNAL	1	21	POTENTIAL.	
FT	CHAIN	22	498	ALPHA-AMYLASE A.	
FT	ACT_SITE	227	227	BY SIMILARITY.	
FT	ACT_SITE	231	231	BY SIMILARITY.	
FT	ACT_SITE	251	251	BY SIMILARITY.	
FT	ACT_SITE	251	251	BY SIMILARITY.	
FT	DISULFID	318	318	BY SIMILARITY.	
FT	DISULFID	51	59	BY SIMILARITY.	
FT	DISULFID	171	185	BY SIMILARITY.	
FT	DISULFID	261	304	BY SIMILARITY.	
FT	DISULFID	461	496	BY SIMILARITY.	
FT	CARBOHYD	218	218	N-LINKED (GLCNAC...) (POTENTIAL).	
SO	SEQUENCE	498 AA;	54880 MW;	7658511BC01A8A01 CRC64;	

Query Match 100.0%; Score 41; DB 1; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGQQLT 8
 |||||
 Db 449 YTAGQQLT 456

RESULT 2

AMTA_ASPOR STANDARD; PRT; 499 AA.
 AC P10529; P11763; Q00250;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
 GN (1,4-alpha-D-glucanohydrolase).
 DE AMY1 AND AMY2 AND AMY3.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 63303;
 RX MEDLINE=89237897; Pubmed=2785629;
 RA Wiersel S., Lachmund A., Wildhardt G., Rutkowski E.;
 RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
 RT intron-exon organization.";
 RL Mol. Microbiol. 3:3-14(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89378767; Pubmed=2789162;
 RA Genes M.J., Dove M.J., Sellig V.L.;
 RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
 RT containing eight introns.";
 RL Gene 79:107-117(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
 RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
 RT of Aspergillus oryzae.";
 RL Agric. Biol. Chem. 53:593-599(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90128276; Pubmed=2612911;
 RA Tsukagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A.,
 RA Ueda S.;
 RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
 RT evidence for multiple related genes.";
 RL Gene 84:319-327(1989).
 RN [5]
 RP SEQUENCE OF 22-499.
 RA Toda H., Kondo K., Narita K.;
 RT "The complete amino acid sequence of Taka-amylase A.";
 RL Proc. Jpn. Acad. B, Phys. Biol. Sci. 58:208-212(1982).
 RN [6]
 RP SEQUENCE OF 206-225.
 RX MEDLINE=74001521; Pubmed=4733850;
 RA Isemura S., Ikenaka T.;
 RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
 RT A with trypsin and chymotrypsin.";
 RL J. Biochem. 74:1-10(1973).
 RN [7]
 RP SEQUENCE OF 433-499.
 RA Narita K.;
 RT Proc. Jpn. Acad. B, Phys. Biol. Sci. 51:285-290(1975).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=80227691; Pubmed=6156152;

RA Matsura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
 RA Toda H., Narita K., Kakudo M.;
 RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
 RT 3-A resolution.";
 RL J. Biochem. 87:1555-1558(1980).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=84212370; Pubmed=6609921;
 RA Matsura Y., Kusunoki M., Harada W., Kakudo M.;
 RT "Structure and possible catalytic residues of Taka-amylase A.";
 RL J. Biochem. 95:697-702(1984).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
 RX MEDLINE=97428212; Pubmed=9283074;
 RA Brozowski A.M., Davies G.J.;
 RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
 RT inhibitor acarbose at 2.0-A resolution.";
 RL Biochemistry 36:10837-10845(1997).
 CC -I CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -I COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -I MISCELLANEOUS: THE SEQUENCE OF AMY1 AND AMY2 IS SHOWN.
 CC -I SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC between the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X12725; CAA31218.1; -;
 DR EMBL: X12726; CAA31219.1; -;
 DR EMBL: X12727; CAA31220.1; -;
 DR EMBL: D00434; BA00036.1; -;
 DR EMBL: M33218; AAA32708.1; -;
 DR PIR: S04548; ALAS1.
 DR PIR: S04549; ALAS3.
 DR PIR: A10627; A10627.
 DR PIR: JK0201; JK0201.
 DR PIR: JS0240; JS0240.
 DR PIR: JT0466; JT0466.
 DR PDB: 2TAA; 15-OCT-89.
 DR PDB: 6TAA; 31-OCT-93.
 DR PDB: 7TAA; 25-NOV-98.
 DR GlycosultedB: P10529; -;
 DR InterPro: IPR000461; Alpha.amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 FT
 FT VARIANT 56 56
 FT VARIANT 172 172
 FT VARIANT 93 94
 FT CONFLICT 106 106
 FT CONFLICT 184 184
 FT CONFLICT 195 195
 FT CONFLICT 255 255
 FT CONFLICT 291 291
 FT CONFLICT 345 345
 N-LINKED (GLCNAC. . .).
 /FTID-CAR.000125.
 O -> R (IN AMY3).
 F -> L (IN AMY3).
 TP -> DC (IN REF. 5).
 O -> T (IN REF. 5).
 D -> Y (IN REF. 3).
 P -> L (IN REF. 3).
 G -> V (IN REF. 3).
 D -> H (IN REF. 4).
 I -> L (IN REF. 5).

```

FT CONFLICT 370 370 L -> A (IN REF. 4).
FT CONFLICT 406 409 WPLY -> PYI (IN REF. 5).
FT CONFLICT 448 448 G -> S (IN REF. 5).
FT CONFLICT 497 497 S -> SD (IN REF. 5 AND 7).
FT HELIX 24 27
FT TURN 28 29
FT STRAND 32 35
FT HELIX 37 40
FT STRAND 42 42
FT TURN 43 44
FT HELIX 53 55
FT STRAND 61 61
FT HELIX 63 67
FT TURN 68 69
FT HELIX 70 74
FT TURN 75 77
FT STRAND 80 83
FT STRAND 87 89
FT STRAND 94 94
FT TURN 95 96
FT STRAND 97 97
FT TURN 100 101
FT STRAND 105 111
FT TURN 113 115
FT TURN 118 119
FT HELIX 120 130
FT TURN 131 132
FT STRAND 134 139
FT STRAND 143 143
FT STRAND 146 146
FT HELIX 150 152
FT TURN 155 156
FT STRAND 159 159
FT HELIX 164 166
FT STRAND 167 167
FT STRAND 172 172
FT HELIX 179 184
FT STRAND 186 188
FT STRAND 193 194
FT STRAND 196 197
FT TURN 199 200
FT HELIX 202 219
FT TURN 220 220
FT STRAND 223 226
FT TURN 227 228
FT HELIX 229 231
FT TURN 234 235
FT HELIX 247 250
FT STRAND 257 265
FT HELIX 266 266
FT STRAND 269 271
FT HELIX 273 283
FT TURN 286 287
FT HELIX 290 303
FT STRAND 307 309
FT TURN 311 312
FT STRAND 317 318
FT HELIX 322 325
FT STRAND 341 341
FT HELIX 345 349
FT TURN 350 351
FT HELIX 352 354
FT TURN 355 355
FT TURN 361 364
FT HELIX 368 371
FT TURN 372 372

```

Query Match 100.0%; Score 41; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8

```

Db 449 YTAGOQLT 456
RESULT 3
ID AMYB_ASPAW STANDARD: PRT: 499 AA.
AC 002906;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alpha-amylase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase B).
GN AMYB.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OVK13F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
RT from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: X52756; CAA36967.1; -.
CC DR HSSP: P10529; YTA.
CC DR InterPro: IPR000461; Alpha-amylase.
CC DR Pfam: PF00128; alpha-amylase; 1.
CC KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
CC Glycoprotein; Signal; Multigene family.
CC FT SIGNAL 1 21
CC FT CHAIN 22 499
CC FT ACT_SITE 227 227
CC FT ACT_SITE 231 231
CC FT ACT_SITE 251 251
CC FT ACT_SITE 318 318
CC FT DISULFID 51 59
CC FT DISULFID 171 185
CC FT DISULFID 261 304
CC FT DISULFID 461 496
CC FT CARBOHYD 218 218
CC SO SEQUENCE 499 AA; 54921 MW; 740B96B11BC01A8A CRC64;

```

Query Match 100.0%; Score 41; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8

Db 449 YTAGOQLT 456

RESULT 4

AMY_ASPSH

```

ID  AMY-ASPSH      STANDARD:      PRT:      499 AA.
AC  P30292;
DT  01-APR-1993 (Rel. 25, Created)
DT  01-APR-1993 (Rel. 25, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE  glucanohydrolase).
GN  AMY.
OS  Aspergillus shirousami.
OC  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC  Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX  NCBI_TaxID=5070;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=92323146; PubMed=1368777;
RX  Shibuya I., Tamura G., Ishikawa T., Hara S.;
RT  "Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its
RT  expression in Saccharomyces cerevisiae."
RL  Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC  -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC  linkages in oligosaccharides and polysaccharides.
CC  -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC  -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: D10461; BAA01255.1; -
DR  PIR: J50663; J50663.
DR  HSP: P10529; 7TAA.
DR  InterPro: IPR000461; Alpha_Amylase.
DR  Pfam: PF00128; alpha-amylase.1.
KW  Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KW  Glycoprotein; Signal.
FT  SIGNAL      1      21
FT  CHAIN       1      499      ALPHA-AMYLASE.
FT  ACT_SITE    227      227
FT  ACT_SITE    231      231
FT  ACT_SITE    251      251      BY SIMILARITY.
FT  ACT_SITE    318      318      BY SIMILARITY.
FT  ACT_SITE    51      59      BY SIMILARITY.
FT  DISULFID    171      185      BY SIMILARITY.
FT  DISULFID    261      304      BY SIMILARITY.
FT  DISULFID    461      496      BY SIMILARITY.
FT  CARBOHYD    218      218      N-LINKED (GLCNAC... ) (BY SIMILARITY).
SQ  SEQUENCE    499 AA: 54852 MW; 1FB7AE50DA01C03F CRC64;

Query Match      100.0%; Score 41; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  1 YTAGOOLT 8
DB  449 YTAGOOLT 456

RESULT  5
NRPF_HAEIN
ID  NRPF_HAEIN      STANDARD:      PRT:      370 AA.
AC  P44942;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Cytochrome c-type biogenesis protein nrpf precursor.
GN  NRPF OR H10934.
OS  Haemophilus influenzae.

```

```

OC  Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC  Haemophilus.
OX  NCBI_TaxID=727;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=RD / KW20 / ATCC 51907;
RX  MEDLINE=95350630; PubMed=7542800;
RA  Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA  Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA  McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA  Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA  Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA  Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA  Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
RA  Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA  Venter J.C.;
RT  "Whole-genome random sequencing and assembly of Haemophilus
RT  influenzae Rd."
RL  Science 269:496-512(1995).
CC  -1- FUNCTION: MAY BE REQUIRED FOR THE BIOGENESIS OF C-TYPE
CC  CYTOCHROMES. POSSIBLE SUBUNIT OF A HEME LYASE (BY SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: Periplasmic (Probable).
CC  -1- SIMILARITY: BELONGS TO THE CCMH/CYCL/CCL2/NRPF FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: U32775; AAC22592.1; ALT_INIT.
DR  TIGR: H10934; -
DR  InterPro: IPR001440; TPR.
DR  Pfam: PF00515; TPR; 1.
KW  Cytochrome c-type biogenesis; Periplasmic; Heme; Signal;
KW  Complete proteome.
FT  SIGNAL      1      21      POTENTIAL.
FT  CHAIN       1      370      CYTOCHROME C-TYPE BIOGENESIS PROTEIN
FT  BINDING     47      47      NRPF.
FT  BINDING     50      50      HEME (COVALENT) (POTENTIAL).
FT  BINDING     50      50      HEME (COVALENT) (POTENTIAL).
SQ  SEQUENCE    370 AA: 42161 MW; 4023B680D7EA9480 CRC64;

Query Match      80.5%; Score 33; DB 1; Length 370;
Best Local Similarity 75.0%; Pred. No. 6.6;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY  1 YTAGOOLT 8
DB  287 YTAGOOLT 294

RESULT  6
VP47_NPYOP
ID  VP47_NPYOP      STANDARD:      PRT:      399 AA.
AC  O10300;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last annotation update)
DE  Viral transcription regulator P47.
GN  P47.
OS  Orgyia pseudotsugata multicausid polyhedrosis virus (OPMPV).
OC  Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC  Nucleopolyhedrovirus.
OX  NCBI_TaxID=164623;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=97271300; PubMed=9126251;
RX  Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RX  Rohrmann G.F.;

```


RT "The sequence of the Orizya pseudotsugata multinucleocapsid nuclear
 RT polyhedrosis virus genome";
 RL Virology 229:381-399(1997).
 CC -1 FUNCTION: IT IS INVOLVED IN REGULATING VIRAL TRANSCRIPTION AT
 CC LATE TIMES POSTINFECTION (BY SIMILARITY).
 CC -1 SUBCELLULAR LOCATION: NUCLEUS OF INFECTED CELLS (BY SIMILARITY).
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: U75930; AAC59044.1;
 DR Transcription regulation; Nuclear protein.
 RW
 SQ SEQUENCE 399 AA; 46357 MW; A5476DDCF83225C0 CRC64;

```

Query Match      78.0%; Score 32; DB 1; Length 399;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 YTAGGQILT 8
    |||
Db 148 YTTGGQQLS 155

```

```

RESULT      7
ID          VP47_NPVAC          STANDARD:          PRT:      401 AA.
AC          P34051:
DT          01-FEB-1994 (Rel. 28, Created)
DT          01-FEB-1994 (Rel. 28, Last sequence update)
DT          01-NOV-1995 (Rel. 32, Last annotation update)
DE          Viral transcription regulator p47.
GN          p47.
OS          Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC          Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC          Nucleopolyhedrovirus.
CX          NCBI_TaxID=46015;
RN          [1]
RP          SEQUENCE FROM N.A.
RC          STRAIN=HR3;
RC          MEDLINE=9323214; PubMed=8474157;
RA          Carstens E.B., Lu A.L., Chan H.L.B.;
RT          "Sequence, transcriptional mapping, and overexpression of p47, a
RL          baculovirus gene regulating late gene expression.";
RL          J. Virol. 67:2513-2520(1993).
RN          [2]
RP          SEQUENCE FROM N.A.
RC          STRAIN=CG;
RC          MEDLINE=94303173; PubMed=8030224;
RA          Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT          "The complete DNA sequence of Autographa californica nuclear
RL          polyhedrosis virus.";
RL          Virology 202:586-605(1994).
CC          -1- FUNCTION: IT IS INVOLVED IN REGULATING VIRAL TRANSCRIPTION AT
CC          LATE TIMES POST INFECTION.
CC          -1- SUBCELLULAR LOCATION: NUCLEUS OF INFECTED CELLS.
CC          -----
CC          THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC          between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC          the European Bioinformatics Institute. There are no restrictions on
CC          use by non-profit institutions as long as its content is in no way
CC          modified and this statement is not removed. Usage by and for commercial
CC          entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC          or send an email to license@isb-sib.ch).
CC          -----
DR          EMBL; L07878; AAA16859.1; -
DR          EMBL; L22858; AAA6670.1; -.
DR          PIR; B45684; B45684.

```

KW	Transcription regulation; Nuclear protein.
FT	VARIANT 131 V -> M (TEMPERATURE SENSITIVE MUTANT).
SD	SEQUENCE 401 AA; 47530 MM; 6E1818027042DC3A CRC64;

```

      78.0%   Score 32; DB 1;   Length 401;
Query Match      Pred. No. 12;
Best local Similarity 75.0%;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 YTAGGQLT 8
      || |||:
Db      144 YTAGGQLS 151

```

RESULT	8
GT21.DROME	STANDARD;
ID	GT21.DROME
AC	09VG98: 09VX91;
DT	16-OCT-2001 (rel. 40, Created)
DT	16-OCT-2001 (rel. 40, Last sequence update)
DT	16-OCT-2001 (rel. 40, Last annotation update)
DE	Glutathione S-transferase D21 (BC 2.5.1.18) (DmgGT21).
GN	GSTD21 OR CG4181.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxId=7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=93252851; PubMed=7683659;
RA	Toung Y.-P., S., Hsieh T.-S., Yu C.-P.D.;
RT	"The glutathione S-transferase D genes. A divergently organized,
RL	intronsless gene family in Drosophila melanogaster.";
RL	J. Biol. Chem. 268: 9737-9746(1993).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BERKELEY;
RX	MEDLINE=20196006; PubMed=10731132;
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Alamatis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Abdill J.F., Agapayani A., An H.-U., Andrews-Plattko C., Baldwin D.,
RA	Balcer R.M., Basu A., Baxendale S., Bayraktaroglu L., Beasley E.M.,
RA	Baesens R.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA	Borkuta D., Botchan M.A., Bouck J., Brockstein P., Brottler P.,
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Cawley S., Dalke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dow I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.
RA	Durbin K.J., Evangelista A.C., Ferraz C., Ferriera S., Fleischmann W.
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA	Hosch D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
RA	Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattel B., McInosh T.C., McLeod M.P., McPherson D.
RA	Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA	Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA	Palazzo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA	Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
RA	Spier E., Sprelding A.C., Stapleton M., Strong R., Sun E.,
RA	Styraskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-T., Wassaman D.A., Weinstock G.M., Weissbach J.,
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

```

RA  Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster."
RL  Science 287:2185-2195(2000)
CC  -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF
CC  EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES (BY
CC  SIMILARITY).
CC  -1- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC  -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. THETA FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: M97702; -; NOT_ANNOTATED_CDS.
DR  FLYBASE: FBgn0010038; GSTd21.
DR  InterPro: IPR004046; GST_C.
DR  Pfam: PF00043; GST_C; 1.
DR  Pfam: PF02798; GST_N; 1.
KW  Transferase; Multigene family.
SQ  SEQUENCE 215 AA; 24535 MW; 221A9961BBF0E5B5 CRC64;

Query Match          75.6%; Score 31; DB 1; Length 215;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1 YTAGOOLT 8
DB  146 YVAGDOLT 153

RESULT  9
GT33_DROME STANDARD; PRT; 215 AA.
ID  GT23_DROME
AC  Q9V9G6; Q9V9G0;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Glutathione S-transferase D23 (EC 2.5.1.18) (DMGST23).
GN  GSTD23 OR CG11512.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxId=7227;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93252851; PubMed=7683659;
RA  Tsoung Y.-P.S., Halen T.-S., Tu C.-P.D.;
RT  "The glutathione S-transferase D genes. A divergently organized,
RT  intronless gene family in Drosophila melanogaster."
RL  J. Biol. Chem. 268:9737-9746(1993).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  STRAIN-BERKELEY;
RA  MEDLINE=20196006; PubMed=10731132.
RA  Adams M.D., Celinker S.E., Holt R.A., Gocayne J.D.,
RA  Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA  George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA  Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,
RA  Brandon R.C., Rogers Y.-H.C., Blazej R.C., Champe M., Pfeiffer B.D.,
RA  Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA  Adair J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA  Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

```

```

RA  Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA  Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA  Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA  de Pablos B., Delcher A., Deng Z., Mays R.D., Dew I., Dietz S.M.,
RA  Doonan K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA  Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA  Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA  Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA  Harris A.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA  Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA  Jaisl M., Kalush F., Kapten G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA  Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA  Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA  Liu X., Mallet Y., McIntosh T.C., McLeod M.P., McPherson D.,
RA  Mekurov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA  Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA  Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacile J.M.,
RA  Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA  Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA  Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA  Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA  Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA  Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA  Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA  Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA  Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster."
RL  Science 287:2185-2195(2000)
CC  -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF
CC  EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES (BY
CC  SIMILARITY).
CC  -1- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC  -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. THETA FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: M97702; -; NOT_ANNOTATED_CDS.
DR  FLYBASE: FBgn0010040; GSTd23.
DR  InterPro: IPR004046; GST_C.
DR  InterPro: IPR004045; GST_N.
DR  Pfam: PF00043; GST_C; 1.
DR  Pfam: PF02798; GST_N; 1.
KW  Transferase; Multigene family.
FT  CONFLICT 142 142 E -> V (IN REF. 1).
SQ  SEQUENCE 215 AA; 24736 MW; 9556E0A1A54BBE1F CRC64;

Query Match          75.6%; Score 31; DB 1; Length 215;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1 YTAGOOLT 8
DB  146 YVAGDOLT 153

RESULT  10
PPIRO_STRCO STANDARD; PRT; 226 AA.
ID  PPIRO_STRCO
AC  Q9RKK6;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Phosphoribosylformylglycinamidine synthase I (EC 6.3.5.3) (PGAM

```

```
DE synthase 1).
GN PURO OR SCD5.14.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetales; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Brown S.P., Harris D., Thomson N.R., Parkhill J., Barrell B.G.,
RA Rajandream M.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + 5'-phosphoribosylformylglycinamide + L-
CC glutamine + H(2)O = ADP + phosphate + 5'-
CC phosphoribosylformylglycinamide + L-glutamate.
CC -1- PATHWAY: DE NOVO PURINE BIOSYNTHESIS; FOURTH STEP.
CC -1- SUBUNIT: HETERODIMER OF TWO SUBUNITS, PURO AND PURL.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity)
CC -1- SIMILARITY: BELONGS TO TYPE-1 GLUTAMINE AMIDOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AL118514; CAB56358.1; -
DR InterPro: IPR000991; GATase.1.
DR PROSITE: PS00442; GATASE_TYPE_I; 1.
FM Purine biosynthesis; Ligase; Glutamine amidotransferase.
KT ACT_SITE 87 87 GATASE (By similarity).
SQ SEQUENCE 226 AA; 24528 MW; 4744E8706CD39D23 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 226;
Best Local Similarity 71.4%; Pred. NO. 11;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOL 7
Db 131 YTAGOEL 137

RESULT 11
ID Y1G9_VIBCH STANDARD: PRT; 282 AA.
AC Q9KKY1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein VCA0969.
GN VCA0969.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RC MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Gill S.R., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Richardson D.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- SIMILARITY: BELONGS TO THE PIRIN FAMILY.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AE004423; AAF96865.1; -
DR TIGR, VCA0969; -
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 282 AA; 31324 MW; 5010C3933BE13197 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 282;
Best Local Similarity 85.7%; Pred. NO. 14;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YAGOOIT 8
Db 226 YAGEOLT 232

RESULT 12
ID NX1B_HUMAN STANDARD: PRT; 442 AA.
AC P58400;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neurexin 1-beta precursor (Neurexin I-beta).
GN NRXN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Yohikawacatarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=Brain;
RP SEQUENCE FROM N.A.
RX MEDLINE=99118874; PubMed=9921901;
RA Kleiderlein J.U., Nilsson P.E., Jessee J., Li W.B., Becker K.G.,
RA Derby M.L., Ross C.A., Margolis R.L.;
RT "CCG repeats in cDNAs from human brain."
RL Hum. Genet. 103:666-673(1998).
RN [2]
RP CONCEPTUAL TRANSLATION OF 1-87.
RA Pardo B.;
RL Unpublished observations (JUN-2001).
CC -1- FUNCTION: NEURONAL CELL SURFACE PROTEIN THAT MAY BE INVOLVED IN
CC CELL RECOGNITION AND CELL ADHESION BY FORMING INTRACELLULAR
CC JUNCTIONS THROUGH BINDING TO NEUROLIGINS. MAY PLAY A ROLE IN
CC FORMATION OR MAINTENANCE OF SYNAPTIC JUNCTIONS. MAY MEDIATE
CC INTRACELLULAR SIGNALING.
CC -1- SUBUNIT: THE CYTOPLASMIC C-TERMINAL REGION BINDS TO CASR. SPECIFIC
CC ISOFORMS BIND NEUROLIGIN, ALPHA-DYSTROGLYCAN AND ALPHA-LATROTOXIN
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein (Potential).
CC -1- ALTERNATIVE PRODUCTS: MULTIPLE ISOFORMS MAY BE PRODUCED BY
CC ALTERNATIVE SPLICING (BY SIMILARITY). BY USE OF ALTERNATIVE
CC PROMOTERS IN THE NEUREXIN 1 GENE, THE ALPHA (AC 90UB1) AND BETA
CC ISOFORMS DIFFER IN THEIR N-TERMINUS.
CC -1- PTM: HIGHLY O-GLYCOSYLATED AND MINOR N-GLYCOSYLATED (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NEUREXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```

CC -----
DR EMBL: AF064842; NOT_ANNOTATED.CDS.
DR MIM: 600565;
DR PROSITE: PS50025; LAM_G-DOMAIN; 1.
KW SIGNAL; Transmembrane; Cell adhesion; Glycoprotein.
FT SIGNAL 1 50
FT CHAIN 51 442
FT DOMAIN 51 363
FT TRANSMEM 364 386
FT DOMAIN 387 442
FT DOMAIN 91 259
FT DOMAIN 289 292
FT DOMAIN 374 377
FT CARBOHYD 188 188
FT CONFLICT 1 53
FT FT
FT FT
SQ SEQUENCE 442 AA; 46627 MW; E365528975198FAD CRC64;
  REF. 1).
  N-LINKED (GLCNAC...) (POTENTIAL).
  MYOIRLCGAEIGSGEGGGGGGGGAGGCRLLAIIVPLTL
  SGLIGVAMGASS -> SRSPSASYEMRGHPV (IN
  REF. 1).

Query Match      75.6%; Score 31; DB 1; Length 442;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGOOLT 8
Db 202 YTAGOOLT 209

RESULT 13
CX1A_PARDE STANDARD; PRT; 554 AA.
AC P08305;
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Cytochrome c oxidase polypeptide I-alpha (EC 1.9.3.1) (Cytochrome AA3
  subunit I-alpha).
GN CTADT OR COT.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S1657;
RA Raitio M., Jalli T., Saraste M.;
RT "Isolation and analysis of the genes for cytochrome c oxidase in
  Paracoccus denitrificans.";
RL EMO J. 6:2825-2833(1987).
  - FUNCTION: SUBUNIT I AND IT FORM THE FUNCTIONAL CORE OF THE ENZYME
  COMPLEX. ELECTRONS ORIGINATING IN CYTOCHROME C ARE TRANSFERRED VIA
  HEME A AND CU(A) TO THE BINUCLEAR CENTER FORMED BY HEME A3 AND
  CU(B). THIS CYTOCHROME C OXIDASE SHOWS PROTON PUMP ACTIVITY
  ACROSS THE MEMBRANE IN ADDITION TO THE ELECTRON TRANSFER.
  - CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
  c + 2 H(2)O.
  - COFACTOR: TWO HEME GROUPS AND COPPER B.
  - PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.
  - SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
  - SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
  - This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb.ch/announce/
  or send an email to license@isb-sib.ch).
  -----
DR EMBL: X05829; CAA29274.1;
DR PIR: S03809; S03809.
DR HSP: P98002; IAR1.

```

```

DR InterPro: IPR000883; COX1.
DR Pfam: PF00115; COX1.1.
DR PRINTS: PR01165; CYCOXIDASE1.
DR PROSITE: PS00077; COX1.1.
KW Oxidoreductase; Heme; Copper; Transmembrane; Inner membrane;
  Respiratory chain; Hydrogen ion transport.
FT METAL 91 91
FT METAL 273 273
FT METAL 277 277
FT METAL 322 322
FT METAL 323 323
FT METAL 406 406
FT METAL 408 408
FT DISULFID 63 77
FT TRANSMEM 26 56
FT TRANSMEM 81 118
FT TRANSMEM 127 148
FT TRANSMEM 175 203
FT TRANSMEM 215 248
FT TRANSMEM 260 295
FT TRANSMEM 301 319
FT TRANSMEM 331 359
FT TRANSMEM 367 390
FT TRANSMEM 399 425
FT TRANSMEM 436 463
FT TRANSMEM 478 508
SQ SEQUENCE 554 AA; 62012 MW; 58AD591FBBDE794 CRC64;

Query Match      75.6%; Score 31; DB 1; Length 554;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YTAGOOLT 8
Db 325 YTAGMSLT 332

RESULT 14
CX1B_PARDE STANDARD; PRT; 558 AA.
AC P98002;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide I-beta (EC 1.9.3.1) (Cytochrome AA3
  subunit I-beta).
GN CTADIT.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PD 1222;
RX MEDLINE=90184495; PubMed=215830;
RA Raitio M., Pispä J.M., Metsu T., Saraste M.;
RT "Are there isoenzymes of cytochrome c oxidase in Paracoccus
  denitrificans?";
RL FEBS Lett. 261:431-435(1990).
  [2]
  X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
  MEDLINE=95379447; PubMed=7651515;
  Iwata S., Ostermeier C., Ludwig B., Michel H.;
  "Structure at 2.8-A resolution of cytochrome c oxidase from
  Paracoccus denitrificans.";
  Nature 376:660-669(1995).
  [3]
  X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
  MEDLINE=98021406; PubMed=9380672;
  Ostermeier C., Harrenga A., Emler U., Michel H.;
  "Structure at 2.7-A resolution of the Paracoccus denitrificans two-
  subunit cytochrome c oxidase complexed with an antibody FV

```

fragment.";
 Proc. Natl. Acad. Sci. U.S.A. 94:10547-10553(1997).
 [4]
 COVALENT BOND.
 MEDLINE=99268331; PubMed=10338009;
 RA Buse G., Soulimane T., Demor M., Meyer H.E., Blueggel M.;
 "Evidence for a copper-coordinated histidine-tyrosine cross-link in
 the active site of cytochrome oxidase.";
 Protein Sci. 8:985-990(1999).
 [5]
 REVIEW.
 MEDLINE=96391941; PubMed=8794157;
 RA Ostermeier C., Iwata S., Michel H.;
 "Cytochrome c oxidase.";
 Curr. Opin. Struct. Biol. 6:460-466(1996).
 CC -I- FUNCTION: SUBUNIT I AND II FORM THE FUNCTIONAL CORE OF THE ENZYME
 COMPLEX. ELECTRONS ORIGINATING IN CYTOCHROME C ARE TRANSFERRED VIA
 HERE A AND CU(A) TO THE BINCLAR CENTER FORMED BY HEME A3 AND
 CU(B). THIS CYTOCHROME C OXIDASE SHOWS PROTON PUMP ACTIVITY
 ACROSS THE MEMBRANE IN ADDITION TO THE ELECTRON TRANSFER.
 CC -I- CATALYTIC ACTIVITY: 4 ferriocytochrome c + O(2) = 4 ferriocytochrome
 c + 2 H(2)O.
 CC -I- COFACTOR: TWO HEME GROUPS AND COPPER B.
 CC -I- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN. Inner membrane.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -I- PTR: HIS-276 AND TYR-280 RESIDUES ARE INVOLVED IN THE FORMATION OF
 A COPPER-COORDINATED COVALENT CROSSLINK AT THE ACTIVE SITE OF THE
 CATALYTIC SUBUNIT I.
 CC -I- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL: Y07533; CAA68821.1; -
 DR PIR: S08270; S08270.
 DR PDB: 1AR1; 1I-FEB-98.
 DR InterPro: IPR000883; COX1.
 DR Pfam: PF00115; COX1; 1.
 DR PRINTS: PR01165; CYCOXIDASE1.
 DR PROSITE: PS00077; COX1; 1.
 KW Oxidoreductase; Heme; Copper; Ion transport; 3D-structure.
 KW Respiratory chain; Hydrogen
 FT DOMAIN 1 28
 FT TRANSMEM 29 59
 FT DOMAIN 60 82
 FT TRANSMEM 83 120
 FT DOMAIN 121 126
 FT TRANSMEM 127 131
 FT DOMAIN 152 176
 FT TRANSMEM 177 206
 FT DOMAIN 207 217
 FT TRANSMEM 218 251
 FT DOMAIN 252 262
 FT TRANSMEM 263 299
 FT DOMAIN 300 303
 FT TRANSMEM 304 331
 FT DOMAIN 332 332
 FT TRANSMEM 333 364
 FT DOMAIN 365 369
 FT TRANSMEM 370 395
 FT DOMAIN 396 404
 FT TRANSMEM 405 437
 FT DOMAIN 438 440
 FT TRANSMEM 441 469
 FT DOMAIN 470 478
 FT TRANSMEM 479 514
 FT DOMAIN 515 558
 FT METAL 94 94
 IRON (HEME A).

FT	METAL	276	276	COPPER B.
FT	METAL	280	280	COPPER B.
FT	METAL	325	325	COPPER B.
FT	METAL	326	326	COPPER B.
FT	METAL	411	411	IRON (HEME A3).
FT	METAL	413	413	IRON (HEME A).
FT	DISULFID	66	80	
FT	SEQUENCE	558 AA;	62439 MW;	AB402453C0C0339E CRC64;

Query Match 75.68; Score 31; DB 1; Length 558;
 Best Local Similarity 75.04; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
 DB 328 YTAGMSLT 335

RESULT 15
 CH11_BACCI STANDARD; PRT; 699 AA.
 ID CH11_BACCI
 AC P20533;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Chitinase A1 precursor (EC 3.2.1.14).
 GN CH11.
 OS Bacillus circulans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1397;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WL-12.
 RA MEDLINE=90368776; PubMed=2203782;
 RX Watanabe T., Suzuki K., Oyanagi W., Ohnishi K., Tanaka H.;
 "Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed
 its evolutionary relationship to Serratia chitinase and to the type
 III homology units of fibronectin.";
 RL J. Biol. Chem. 265:15659-15665(1990).
 RN [2]
 RP MUTAGENESIS.
 RC STRAIN-WL-12.
 RA MEDLINE=93366760; PubMed=8103047;
 RX Watanabe T., Kohori K., Miyashita K., Fujii T., Sakai H.,
 Uchida M., Tanaka H.;
 "Identification of glutamic acid 204 and aspartic acid 200 in
 chitinase A1 of Bacillus circulans WL-12 as essential residues for
 chitinase activity.";
 RL J. Biol. Chem. 268:18567-18572(1993).
 CC -I- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
 acetyl-D-glucosamine polymers of chitin.
 CC -I- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -I- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 HYDROLASES).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL: M57601; AAA81528.1; -
 DR PIR: A38368; A38368.
 DR HSSP: P07254; 1CPN.
 DR InterPro: IPR003610; Chitin_bin3.
 DR InterPro: IPR001579; Chitinase_2.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003962; FNIII_repeat.
 DR InterPro: IPR001223; Glyco_hydro_18.

DR Pfam; PF02839; CBD_7; 1.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PRINTS; PR00014; FNtypeIII.
DR SMART; SM00495; ChlBD3; 1.
DR SMART; SM00060; FN3; 2.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Signal; Repeat.
FT SIGNAL 1 41
FT CHAIN 42 699 CHITINASE A1.
FT DOMAIN 42 460 CATALYTIC.
FT DOMAIN 465 549 FIBRONECTIN TYPE-III (R-1).
FT DOMAIN 560 644 FIBRONECTIN TYPE-III (R-2).
FT ACT_SITE 204 204 PROTON DONOR (PROBABLE).
FT MUTAGEN 200 200 D->N: DECREASE IN ACTIVITY.
FT MUTAGEN 200 200 D->E: NO CHANGE IN ACTIVITY.
FT MUTAGEN 204 204 E->D,Q: LOSS OF ACTIVITY.
SQ SEQUENCE 699 AA; 73677 MW; AC7C9B22E2987643 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 699;

Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGQLT 8
| | | | | : |
Db 662 YTAGQLVT 669

Search completed: June 13, 2002, 09:11:55
Job time: 1035 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:55:55 ; Search time 73.31 Seconds
(Without alignments)
10.486 Million cell updates/sec

Title: US-09-710-339-2_COPY_448_455
Perfect score: 41
Sequence: 1 YTAGOOLR 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	41	100.0	478	2 JK0201	alpha-amylase (EC
2	41	100.0	498	2 A48305	alpha-amylase (EC
3	41	100.0	499	1 A1AS1	alpha-amylase (EC
4	41	100.0	499	1 A1AS3	alpha-amylase (EC
5	41	100.0	499	2 B48305	alpha-amylase (EC
6	41	100.0	499	2 JS0663	alpha-amylase (EC
7	41	100.0	499	2 JT0466	alpha-amylase (EC
8	41	100.0	499	2 JN0588	alpha-amylase (EC
9	34	82.9	849	2 D82510	chitinase VCA0027
10	33	80.5	384	2 H64161	hypothetical prote
11	32	78.0	390	2 T30395	probable transcrip
12	32	78.0	399	2 T41786	p47 orf40 - Bombyx
13	32	78.0	399	2 T10314	viral transcriptio
14	32	78.0	401	2 H72854	viral transcriptio
15	32	78.0	1250	2 T27706	hypothetical prote
16	31	75.6	214	2 D46681	glutathione transf
17	31	75.6	214	2 E46681	hypothetical prote
18	31	75.6	221	2 F81876	hypothetical prote
19	31	75.6	242	2 T13190	hypothetical prote
20	31	75.6	282	2 H82394	hypothetical prote
21	31	75.6	347	2 D83653	oligopeptide ABC t
22	31	75.6	554	2 S03809	cytochrome-c oxida
23	31	75.6	558	2 S08270	cytochrome-c oxida
24	31	75.6	699	2 A38368	chitinase (EC 3.2.
25	30	75.6	3020	2 A43932	mucin 2 precursor,
26	30	73.2	104	2 T46246	hypothetical prote
27	30	73.2	108	2 E72766	hypothetical prote
28	30	73.2	255	2 D57143	penicillin-binding
29	30	73.2	283	2 F87323	hypothetical prote

30	30	73.2	289	2 A10446	probable carbon-ni
31	30	73.2	331	2 H83240	probable ATP-bindi
32	30	73.2	339	2 AF0300	probable oxidoredu
33	30	73.2	388	2 A75318	conserved hypothet
34	30	73.2	410	2 T36610	hypothetical prote
35	30	73.2	425	2 E71075	hypothetical prote
36	30	73.2	449	2 A24993	probable proton gl
37	30	73.2	452	2 H84772	cellulase (EC 3.2.
38	30	73.2	474	2 S41117	probable serine ca
39	30	73.2	505	2 AC1469	nitric-oxide reduc
40	30	73.2	507	2 T35677	intercalin like pr
41	30	73.2	609	2 D87049	probable hydrolase
42	29	70.7	53	2 S78004	hypothetical prote
43	29	70.7	160	2 F70049	cuticle structural
44	29	70.7	191	2 D82425	flagellin homolog
45	29	70.7	211	2 E83116	guanylate cyclase-50S ribosomal prot

ALIGNMENTS

RESULT 1

JK0201
alpha-amylase (EC 3.2.1.1) - Aspergillus oryzae
N:Alternate names: glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: JK0201

R:Toda, H.; Kondo, K.; Narita, K.
Proc. Jpn. Acad. 58B, 208-212, 1982
A:Title: The complete amino acid sequence of Taka-amylase A.
A:Reference number: JK0201

A:Accession: JK0201
A:Molecule type: protein
A:Residues: 1-478 <TOD>
C:Comment: One atom of calcium per molecule is essential for the activity.
C:Comment: This enzyme is a glycoprotein.
C:Comment: See also PIR:JT0466 and PIR:J50240.
C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
F:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:173-300/Domain: alpha-amylase core homology <AMY>
F:197/Binding site: carbohydrate (asn) (covalent) #status experimental
F:210/230/297/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 41; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOOLR 8
DB 427 YTAGOOLR 434

RESULT 2
A48305
alpha-amylase (EC 3.2.1.1) A precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: A48305

R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylase genes from As
A:Reference number: A48305; MUID:90254827
A:Accession: A48305

A:Molecule type: DNA
A:Status: preliminary; not compared with conceptual translation
A:Residues: 1-498 <KOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 41; DB 2; Length 498;
 Best Local Similarity 100.0%; Pred. No. 0.46;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
 |||||
 Db 449 YTAGOQLT 456

RESULT 3
 ALAS1
 alpha-amylase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae

N:Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 15-Sep-2000
 C:Accession: S04548; A33214; J50240; A91930; A93767; A10627

R:Witsel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon
 A:Reference number: S04548; MUID:89237897

A:Accession: S04548
 A:Molecule type: DNA

A:Residues: 1-499 <MIR>
 A:Cross-references: EMBL:X12725; NID:g2430; PIDN:CAA11218.1; PID:g295921

A:Genetics: AMY1
 A:Accession: A33214

A:Molecule type: mRNA
 A:Residues: 1-499 <M12>

A:Cross-references: GB:X12725; NID:g2430; PIDN:CAA11218.1; PID:g295921
 R:Gene, M.J.; Dove, M.J.; Seligy, V.L.

Gene 79, 107-117, 1989
 A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing

A:Reference number: J50240; MUID:89378767
 A:Accession: J50240

A:Molecule type: DNA
 A:Residues: 1-499 <GEN>

A:Genetics: AMY2
 A>Note: the authors refer to this as isozyme II

R:Isemura, S.; Ikenaka, T.
 J. Biochem. 74, 1-10, 1973

A:Reference number: A91930; MUID:74001521
 A:Accession: A91930

A:Molecule type: protein
 A:Residues: 206-225 <ISE>

R:Marita, K.
 Proc. Jpn. Acad. 51, 285-290, 1975

A:Reference number: A93767
 A:Accession: A93767

A:Molecule type: protein
 A:Residues: 434-443, 446-447, 'Q', 449-458, 'GTRV', 459-464, 467-468, 'B', 470, 'B', 472-499 <NAR>

J. Biochem. 95, 697-702, 1984
 A:Title: Structure and possible catalytic residues of Taka-amylase A.

A:Reference number: A37454; MUID:84212370
 C:Contents: annotation; X-ray crystallography, 3.0 angstroms

R:Swift, H.J.; Brady, L.; Derenda, Z.S.; Dodson, E.J.; Turkenburg, J.P.; Wilkinson, A.
 submitted to the Brookhaven Protein Data Bank, August 1992

A:Reference number: A51548; PDB:6FRA
 C:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 22-497

C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics: <AMY>

A:Gene: amy1
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Genetics: <AMY>

A:Gene: amy2; AmyII
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; po
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-499/Product: alpha-amylase 1 #status experimental <MAT>
 F:194-321/Domain: alpha-amylase core homology <AMY>
 F:51-59,171-185,261-304,461-496/disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 41; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 0.46;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
 |||||
 Db 449 YTAGOQLT 456

RESULT 4
 ALAS3
 alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae

N:Alternate names: alpha-amylase isozyme I; glycogenase; Taka-amylase A
 C:Species: Aspergillus oryzae

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
 C:Accession: S04549; A33215; A44713

R:Witsel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989

A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exo
 A:Reference number: S04548; MUID:89237897

A:Accession: S04549
 A:Molecule type: DNA

A:Residues: 1-499 <MIR>
 A:Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA11220.1; PID:g295922

A:Accession: A33215
 A>Status: translation not shown

A:Molecule type: mRNA
 A:Residues: 1-499 <M12>

A:Cross-references: GB:X12727; NID:g2454; PIDN:CAA11220.1; PID:g295922
 R:Gene, M.J.; Dove, M.J.; Seligy, V.L.

Gene 79, 107-117, 1989
 A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each contain

A:Reference number: J50240; MUID:89378767
 A:Accession: A44713

A:Molecule type: DNA
 A:Residues: 1-499 <GEN>

A>Note: the authors refer to this as isozyme I
 R:Matsura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.

J. Biochem. 95, 697-702, 1984
 A:Title: Structure and possible catalytic residues of Taka-amylase A.

A:Reference number: A37454; MUID:84212370
 C:Contents: annotation; X-ray crystallography, 3.0 angstroms

C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics: amy1

A:Gene: amy3; AmyI
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3

C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology

C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me
 F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-499/Product: alpha-amylase 3 #status experimental <MAT>
 F:194-321/Domain: alpha-amylase core homology <AMY>

F:51-59,171-185,261-304,461-496/disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 41; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
|||||
DB 449 YTAGOQLT 456

RESULT 5
B48305
alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: B48305
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Asper
A:Reference number: A48305; MUID:90234827
A:Accession: B48305
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499 <NOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>
F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 41; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
|||||
DB 449 YTAGOQLT 456

RESULT 6
J50663
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
C:Species: Aspergillus sp.
C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
C:Accession: J50663
R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
Biotech. Biotechnol. Biochem. 56, 174-179, 1992
A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shirousamii and its expression
A:Reference number: J50663; MUID:9232146
A:Accession: J50663
A:Molecule type: mRNA
A:Residues: 1-499 <SHI>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <ALP>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 41; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
|||||
DB 449 YTAGOQLT 456

RESULT 7
JT0466
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: glycosidase; Taka-amylase A
C:Species: Aspergillus oryzae
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: JT0466
R:Tada, S.; Iimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
Agric. Biol. Chem. 53, 593-599, 1989
A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Asperg
A:Reference number: JT0466
A:Accession: JT0466
A:Molecule type: DNA
A:Residues: 1-499 <TAD>
C:Comment: See also PIR:JK0201 and PIR:J50240.
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics:
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <AMY>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 41; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
|||||
DB 449 YTAGOQLT 456

RESULT 8
JN0588
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: Taka-amylase A
C:Species: Aspergillus oryzae
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
C:Accession: JN0588
R:Tsukagoshi, N.; Fukukawa, M.; Nagaba, H.; Kirita, N.; Tsudoi, A.; Udaka, S.
Gene 84, 319-327, 1989
A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for
A:Reference number: JN0588; MUID:90128276
A:Accession: JN0588
A:Molecule type: mRNA
A:Residues: 1-499 <TSU>
C:Comment: The alpha amylases are encoded by multigene family.
C:Genetics:
A:Gene: Taa-G1
A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <AMY>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 41; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
|||||
DB 449 YTAGOQLT 456

Db 449 YTAGOQLT 456
|||||||
RESULT 9
D82510
Chinese VCA0027 [Imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82510
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, H.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: D82510
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-849 <HEI>
A:Cross-references: GB:AE004346; GB:AE003853; NID:96657401; PIDN:AAF95941.1; GSPDB:GN001
C:Genetics:
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
A:Gene: VCA0027
A:Map position: 2
Query Match 82.9%; Score 34; DB 2; Length 849;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;
Indels 0;
OY 1 YTAGOQLT 8
|||||
Db 805 YTGGOQVT 812
RESULT 10
H64161
Hypothetical protein HI0934 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
C:Accession: H64161
R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630
A:Accession: H64161
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-384 <TTGR>
A:Cross-references: GB:U32775; GB:U42023; NID:91573951; PIDN:AAC22592.1; PID:91573955; T
Query Match 80.5%; Score 33; DB 2; Length 384;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 YTAGOQLT 8
|||||
Db 301 YACGOQMT 308
RESULT 11
T30395
probable transcription regulator p47 - Lymantria dispar nuclear polyhedrosis virus
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
C:Accession: T30395
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohl
Virology 253, 17-34, 1999

A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria
A:Reference number: Z20836; MUID:99124785
A:Accession: T30395
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-390 <KU2>
A:Cross-references: EMBL:AF081810; NID:93822234; PIDN:AAC70233.1; PID:93822282
C:Keywords: transcription factor
Query Match 78.0%; Score 32; DB 2; Length 390;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 YTAGOQLT 8
|||||
Db 140 YTLGOQLS 147
RESULT 12
T41786
P47 orf40 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C:Species: Bombyx mori nuclear polyhedrosis virus, BMSNPV
A:Variety: Isolate T3
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T41786
R:Gomi, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A:Reference number: Z22020; MUID:99281911
A:Accession: T41786
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-399 <KAM>
A:Cross-references: EMBL:L33180; NID:93745835; PIDN:AAC63715.1; PID:93745868
A:Experimental source: isolate T3
C:Genetics:
A:Note: p47
Query Match 78.0%; Score 32; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 YTAGOQLT 8
|||||
Db 144 YTLGOQLS 151
RESULT 13
T10314
viral transcription regulator p47 - Orgyia pseudotsugata nuclear polyhedrosis virus
C:Species: Orgyia pseudotsugata nuclear polyhedrosis virus, OPMNPV
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10314
R:Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohrmann, G.F.
Virology 229, 381-399, 1997
A:Title: The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear polyhedro
A:Reference number: Z17011; MUID:97271300
A:Accession: T10314
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-399 <AHR>
A:Cross-references: EMBL:U75930; NID:92934903; PIDN:AAC59044.1; PID:91911291
C:Keywords: transcription regulation
Query Match 78.0%; Score 32; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 YTAGOQLT 8
|||||

Db 148 YTLGQQLS 155

RESULT 14

H72854

Viral transcription regulator p47 - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV

A:Note: dsDNA virus

C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 21-Jul-2000

C:Accession: H72854; B45684

R:Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.

Virology 202, 586-605, 1994

A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.

A:Reference number: A72850; MUID:94303173

A:Accession: H72854

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-401 <AYR>

A:Cross-references: GB:L22858; NID:g510708; PIDN:AAA66670.1; PID:g559109

R:Garstens, E.B.; Lu, A.L.; Chan, H.L.

J. Virol. 67, 2513-2520, 1993

A:Title: Sequence, transcriptional mapping, and overexpression of p47, a baculovirus gen

A:Reference number: A45684; MUID:93233214

A:Accession: B45684

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-401 <CAK>

A:Cross-references: GB:L07878; NID:g293990; PIDN:AAA16859.1; PID:g293992

A:Note: sequence extracted from NCBI backbone (NCBIN:129587, NCBIF:129589)

C:Genetics:

A:Gene: Ac-p47

C:Keywords: nucleus; transcription regulation

Query Match 78.0%; Score 32; DB 2; Length 401;

Best Local Similarity 75.0%; Pred. No. 36;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8

Db 144 YTLGQQLS 151

RESULT 15

T27706

hypothetical protein ZK1193.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27706

R:Geisler, C.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid ZK1193.

A:Reference number: Z20409

A:Accession: T27706

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1250 <GEI>

A:Cross-references: EMBL:U41553; PIDN:AAA83291.1; CESP:ZK1193.2

C:Genetics:

A:Gene: CESP:ZK1193.2

A:Introns: 95/2; 228/1; 302/1; 355/2; 388/1; 400/1; 451/1; 514/1; 593/1; 643/3; 822/2; 8

Query Match

Best Local Similarity 78.0%; Score 32; DB 2; Length 1250;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YTAGOQ 6

Db 698 YTAGOQ 703

Search completed: June 13, 2002, 08:55:56
Job time: 376 sec

Fri Jun 14 10:36:50 2002

us-09-710-339-2_copy_448_455.rpr

Page 6

PT liquefaction agents
XX
PS Claim 1; Page 7; 20pp; English.
XX
XX The sequence as that of the *Aspergillus oryzae* alpha amylase, sold
CC commercially as FRINGAMYL (TM) by Novo Nordisk A/S. The sequence can
CC be mutated by substitution of one or more of its methionine residues
CC for any amino acid other than cysteine. The mutant alpha amylase
CC exhibits a better activity level and better stability in the
CC presence of oxidizing agents than previous mutant alpha amylases,
CC and improved thermostability at moderately low pH. The enzyme can
CC be used as an additive for detergents, dishwashing agents and
CC liquefaction agents.
XX
SQ Sequence 478 AA;

OY 1 YTAGOQLT 8
 |||||
Db 428 ytaagqlt 435

Query Match 100.0%; Score 41; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.78; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

RESULT 2
AAR79025
ID AAR79025 standard; protein; 478 AA.
XX
AC AAR79025;
XX
DT 22-MAR-1996 (first entry)
XX
DE Mature taka-amylase A.
XX
KW Wild type; neopullulanase; B. stearotherophilus; mutant;
KW food industry; modification; hydrophobicity; replacement
KW insertion; deletion.
XX
OS Aspergillus oryzae.
XX
FH Key location/Qualifiers
FH Disulfide-bond 30..38
FT Disulfide-bond 150..164
FT Disulfide-bond 240..283
FT Disulfide-bond 439..474
XX
PN JP07177891-A.
XX
PD 18-JUL-1995.
XX
PF 31-OCT-1994; 94JP-0288658.
XX
PR 12-NOV-1993; 93JP-0306096.
XX
PA (EZAK) EZAKI GLICO CO.
PA (NIDE) NEC CORP.
XX
DR WPI; 1995-279919/37.
XX
PT Modifying a transferase by enhancing hydrophobicity of a selected
PT site increases transfer activity, also new mutant
PT neo-pullulanase(s)
XX
PS Disclosure; Page 10-11; 18pp; English.
XX
CC This sequence represents the mature form of taka-amylase from *A. oryzae*.
CC This sequence was used in a method for the generation of mutant
CC pullulanases for use in the food industry (see also AAR79026-28). The
CC wild type pullulanase enzyme was modified by the method of the
CC invention for enhancing the hydrophobicity of a selected site of the
CC pullulanase. The method comprises replacement of a group in the

CC selected site with a hydrophobic group, replacement of an amino acid
CC with a hydrophobic amino acid, and/or insertion or deletion of a
CC hydrophobic amino acid from the selected site. The method was used
CC to produce neopullulanases Y377E, S422V and M375L.
XX
SQ Sequence 478 AA;

OY 1 YTAGOQLT 8
 |||||
Db 427 ytaagqlt 434

RESULT 3
AAR72450
ID AAR72450 standard; Protein; 478 AA.
XX
AC AAR72450;
XX
DT 01-DEC-1995 (first entry)
XX
DE Aspergillus oryzae alpha amylase (mature protein).
XX
KW Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
KW washing; Bacillus licheniformis; Bacillus amyloliquefaciens;
KW Bacillus stearotherophilus; dyeing; bleaching; scouring; textile;
KW thermostable.
XX
OS Aspergillus oryzae.
XX
PN WO9510603-A.
XX
PD 20-APR-1995.
XX
PF 05-OCT-1994; 94WO-DK00370.
XX
PR 08-OCT-1993; 93DK-0001133.
PR 02-FEB-1994; 94DK-0000140.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Svendsen A, Thellersen M;
PI Van der zee P, Bisgardfrantzen H, Borchert T;
XX
DR WPI; 1995-161790/21.
XX
PT New Bacillus derived alpha-amylase variants - having amino acid
PT modifications to improve washing and/or dishwashing performance
XX
PS Disclosure; Page 75-76; 105pp; English.
XX
CC Variant alpha amylase enzymes which have improved washing and/or
CC as detergent additives. The enzymes have one or more amino acid
CC residues added, deleted or substituted. The variants can also be
CC used for textile desizing prior to scouring, bleaching and dyeing.
CC The variants have improved thermostability, acid/alkaline stability,
CC low temperature optimum, pH optimum; higher hydrolysis velocity and
CC improved tolerance to other composition constituents, e.g. oxidation
CC agents.
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 41; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.78; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

OY 1 YTAGOQLT 8
 |||||

```

Db      428 ytagqqlt 435

RESULT      4
ID      AAR78270
AA      AAR78270 standard; Protein; 478 AA.
XX
XX      AAR78270;
AC
DT      17-JAN-1996 (first entry)
XX
XX      Aspergillus oryzae alpha amylase (mature protein).
DE
XX      Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;
KM      starch; thermotable; methionine; Bacillus licheniformis;
KW      Bacillus amyloliquefaciens; Bacillus stearotheophilus;
KM      Aspergillus oryzae.
XX
XX      Aspergillus oryzae.
OS
PN      WO9521247-A1.
XX
XX      10-AUG-1995.
PD
XX      05-OCT-1994, 94WO-DK00371.
PF
XX      02-FEB-1994; 94DK-0000141.
PR
XX      (NOVO ) NOVO-NORDISK AS.
PA
PI      Marcker D, Nilsson TE, Pedersen HH, Toft AH;
XX      WPI: 1995-283767/37.
DR
XX
XX      Use of an oxidation stable alpha-amylase - for simultaneous desizing
PT      and bleaching or scouring of fabrics contg. starch or starch derivs.
PS
XX      Disclosure; Page 25-26; 37pp; English.
XX
XX      Oxidation stable alpha amylases can be used for the simultaneous
CC      desizing and bleaching or scouring of a fabric comprising starch or
CC      starch derivatives. They exhibit a better heat stability,
CC      especially in the presence of oxidizing agents. They are obtained
CC      from a parent alpha amylase by replacing one or more methionine
CC      residues with any amino acid different from Cys or Met, preferably
CC      Leu, Thr, Ala, Gly, Ser, Ile or Asp. The parent alpha amylase is
CC      pref. derived from a Bacillus species, although alpha amylases of
CC      fungal origin can also be used. This sequence is the wild type
CC      (unmodified) alpha amylase of Aspergillus oryzae.
XX
XX      Sequence 478 AA:
SQ

Query Match      100.0%; Score 41; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0.

Oy      1 YTAGQQLT 8
        |||||||
Db      428 ytagqqlt 435

RESULT      5
ID      AAM14500
AA      AAM14500 standard; protein; 478 AA.
XX
XX      AAM14500;
AC
DT      04-JUN-1997 (first entry)
XX
XX      Aspergillus oryzae alpha-amylase (mature protein).
DE
KW      alpha-amylase; mature; Bacillus licheniformis; termamyl; fungamyl;

```

XX	Aspergillus oryzae.	
XX		
EH	Key	Location/Qualifiers
EH	Misc-difference 13..45	/label= loop 1 modification region
EH		/note= "at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to 7-23 of AAM14499 is deleted or replaced with a fragment corresponding to this fragment; claim 33"
EH	Misc-difference 14..40	/label= loop 1 modification region
EH		/note= "preferred region where at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to 8-18 of AAM14499 is deleted or replaced with a fragment corresponding to this fragment; claim 30"
EH	Misc-difference 28..42	/label= loop 1 modification region
EH		/note= "at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to 12-19 of AAM14499 is deleted or replaced with a fragment corresponding to this fragment; claim 30"
EH	Misc-difference 32..38	/label= loop 1 modification region
EH		/note= "preferred region where at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to 14-15 of AAM14499 is deleted or replaced with a fragment corresponding to this fragment; claim 18"
EH	Misc-difference 66..84	/label= loop 2 modification region
EH		/note= "at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to 44-57 of AAM14499 is deleted or replaced with a fragment corresponding to this fragment; claim 18"
EH	Misc-difference 70..78	/label= loop 2 modification region
EH		/note= "preferred region where at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to 48-51 of AAM14499 is deleted or replaced with a fragment corresponding to this fragment; claim 24"
EH	Misc-difference 98..120	/label= loop 3 modification region
EH		/note= "at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to 117-185 of AAM14499 is deleted or replaced with a fragment corresponding to this fragment; claim 24"
EH	Misc-difference 102..206	/label= loop 3 modification region
EH		/note= "preferred region where at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to 121-189 of AAM14499 is deleted or replaced with a fragment corresponding to this fragment; claim 24"
EH	Misc-difference 121..174	/note= "at least one amino acid residue of a parent alpha-amylase (used as a template for a variant) corresponding to this fragment is deleted or replaced with a fragment corresponding to 102-206 of AAM14499; claim 41"
EH		/note= "preferred region where at least one amino acid

```

FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to this
FT fragment is deleted or replaced with a fragment
FT corresponding to 102-199 of AAM14499; claim 42"
FT
FT Misc-difference 165..177
FT /label= loop 3 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 195-202 of AAM14499 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 21"
FT
FT Misc-difference 166..173
FT /label= loop 3 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 196-198
FT of AAM14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 23"
FT Misc-difference 181..184
FT /note= "an amino acid fragment corresponding to this
FT region is deleted from the parent sequence of a
FT variant Fungamyl; claim 43"
FT
FT Misc-difference 291..313
FT /label= loop 8 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 322-346 of AAM14498 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 36"
FT
FT Misc-difference 297..313
FT /label= loop 8 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 325-345
FT of AAM14498 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 38"
FT
XX WO9623874-A1.
XX
XX 08-AUG-1996.
XX
XX 05-FEB-1996; 96WO-DK00057.
XX
XX 10-NOV-1995; 95DK-0001256.
XX 03-FEB-1995; 95DK-0000128.
XX 23-OCT-1995; 95DK-0001192.
XX
XX (NOVO ) NOVO-NORDISK AS.
XX
XX Bisgard-frantzen H, Borchert TV, Svendsen A;
XX WPI; 1996-371424/37.
XX
XX Alpha-amylase variants and methods of production - have altered
XX PT properties such as calcium dependency, substrate binding and
XX PT stability
XX
XX disclosure: Page 87-88; 171pp; English.
XX
XX The present sequence is the mature Aspergillus oryzae alpha-amylase (A).
XX CC Variants of parent termamyl- and fungamyl-like alpha-amylases (and
XX CC methods of constructing them) are claimed. Examples of variants are
XX CC featured above. The variants have altered properties such as calcium
XX CC dependency, substrate binding and stability. Also one or more proline or
XX CC cysteine residues in the variant is modified or replaced with a
XX CC non-proline or non-cystein residue such as alanine. The variants can be
XX CC used for (dish)washing, as detergent additives or for fabric desizing or
XX CC starch liquefaction. They can also be used for the production of
XX CC sweeteners and ethanol from starch. See also AAM14498-99.
XX
XX Sequence 478 AA:
SQ

```

```

OY 1 YTAGQOLT 8
Db 428 ytagqolt 435
Query Match 100.0%; Score 41; DB 17; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
AAB84206
ID AAB84206 standard; Protein: 498 AA.
XX
AC AAB84206;
XX
XX 06-AUG-2001 (first entry)
XX
XX Amino acid sequence of a fungamyl-like alpha-amylase.
DE
XX Fungamyl-like alpha-amylase; glucoamylase; dextrinisation; maltose;
XX KW alcohol; starch; dough improver; brewing; starch liquefaction.
XX
XX Aspergillus oryzae.
XX OS
XX WO200134784-A1.
XX PN
XX 17-MAY-2001.
XX PD
XX 10-NOV-2000; 2000WO-DK00626.
XX PF
XX 10-NOV-1999; 99DK-0001617.
XX PR
XX (NOVO ) NOVOZYMES AS.
XX
XX Bisgard-Frantzen H, Svendsen A, Pedersen S;
XX PI
XX WPI; 2001-367478/38.
XX DR
XX N-PSDB: AAF90208.
XX
XX New variant of Fungamyl-like alpha-amylase, useful for production of
XX PT maltose syrups, includes mutations that improve stability against heat
XX PT and acidic pH -
XX
XX Claim 1; Page 42-45; 49pp; English.
XX PS
XX
XX The present sequence represents a fungamyl-like alpha-amylase. The
XX CC specification describes variants of this fungamyl-like alpha-amylase,
XX CC which have an alteration in one the amino acid regions 98-110, 150-160,
XX CC 161-167, 280-288, 448-455, and 468-473. Each alteration is a deletion
XX CC or substitution of an amino acid or an insertion of an amino acid
XX CC downstream of a particular position. The variants retain alpha-amylase
XX CC activity, and have better heat stability and/or stability at acidic pH,
XX CC relative to wild-type enzyme. The variants can therefore be used at
XX CC higher temperatures (more efficient conversion or faster reaction, and
XX CC have reduced need for cooling and reduced risk of contamination). The
XX CC variants may also be used in conjunction with other enzymes,
XX CC particularly glucoamylase during dextrinisation. The variants are
XX CC used to produce syrups, particularly of high maltose content, or alcohol,
XX CC from starch; as dough improver for baked goods; in brewing, to increase
XX CC fermentability of the wort; and for liquefaction of starch.
XX
XX Sequence 498 AA:
SQ
Query Match 100.0%; Score 41; DB 22; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 YTAGQOLT 8
Db 448 ytagqolt 455

```


RESULT 7
 ABG20974
 ID ABG20974 standard; Protein: 1626 AA.
 XX
 AC ABG20974:
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #20965.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 OS
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001MO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEC INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 P1
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS85161.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID NO 51333; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1626 AA;

Query Match 78.0%; Score 32; DB 22; Length 1626;
 Best Local Similarity 85.7%; Pred. No. 3.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YTAGOQL 7
 |||||:
 DB 165 ytagkql 171

RESULT 8
 AAG64426
 ID AAG64426 standard; Peptide: 52 AA.
 XX
 AC AAG64426:
 XX
 DT 19-SEP-2001 (first entry)
 XX
 DE B. circulans chitin-cellulose binding domain (CBD) peptide.
 XX
 KW Heteromer peptide; chitin cellulose binding domain; CBD;
 KW protein immobilisation.
 XX
 OS Bacillus circulans.
 OS
 PN W0200131038-A1.
 XX
 PD 03-MAY-2001.
 XX
 PF 19-OCT-2000; 2000MO-JP07275.
 XX
 PR 22-OCT-1999; 99JP-0301699.
 XX
 PA (FUJI) FUJISAWA PHARM CO LTD.
 XX
 PI Tanaka A, Ueda M, Nagao K;
 XX
 DR WPI; 2001-417561/44.
 DR
 XX
 PT Heteromer peptides produced by auto-splicing of a precursor comprise
 PT subunits with chitin cellulose binding domain
 XX
 PS Disclosure; Page 85-86; 100pp; Japanese.
 XX
 CC The present sequence is that of the Bacillus circulans chitin-cellulose
 CC binding domain (CBD) peptide of the invention. The invention relates to
 CC heteromer peptides, produced by cleaving precursor peptides, comprise at
 CC least one subunit with an added chitin cellulose binding domain (CBD)
 CC useful for immobilising proteins for industrial use. The CBD fusion
 CC allows immobilisation of proteins without loss of activity and
 CC immobilised heteromer peptides may be produced with high yield.
 XX
 SQ Sequence 52 AA;

Query Match 75.6%; Score 31; DB 22; Length 52;
 Best Local Similarity 75.0%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YTAGOQLT 8
 |||||:
 DB 15 ytagqlvt 22

RESULT 9
 ABB65782
 ID ABB65782 standard; Protein: 215 AA.
 XX
 AC ABB65782:
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 24138.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 OS
 PN W0200171042-A2.
 PN
 XX
 PD 27-SEP-2001.
 XX

```

PF 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
DR WPI: 2001-656860/75.
DR N-PSDB; ABL09885.
XX
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosure; SEQ ID NO 24138; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (AB157737-AB12072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 215 AA:
SQ

Query Match 75.6%; Score 31; DB 22; Length 215;
Best Local Similarity 75.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
   1 11 111
Db 146 yvagsqlt 153

RESULT 10
AB171287
ID ABB71287 standard; Protein: 215 AA.
XX
XX ABB71287;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 40653.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI: 2001-656860/75.
DR N-PSDB; ABL15390.
XX

```

```

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 40653; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (AB157737-AB12072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 215 AA:
SQ

Query Match 75.6%; Score 31; DB 22; Length 215;
Best Local Similarity 75.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
   1 11 111
Db 146 yvagsqlt 153

RESULT 11
ABB66336
ID ABB66336 standard; Protein: 223 AA.
XX
XX ABB66336;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 23800.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI: 2001-656860/75.
DR N-PSDB; ABL10439.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosure; SEQ ID NO 25800; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX

```

CC sequences (AB101840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 223 AA;

Query Match 75.6%; Score 31; DB 22; Length 223;
Best Local Similarity 75.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
|||
Db 151 yvagnqtl 158

RESULT 12

ABBS9066
ID ABB59066 standard; Protein; 302 AA.

AC ABB59066;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 3990.

KW Drosophila: developmental biology; cell signalling; insecticide;
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

XX N-PSDB; AB103169.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

PS Disclosure: SEQ ID NO 3990; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-ABL16175) and the encoded proteins
CC sequences (ABB57737-ABB72072).
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 302 AA;

Query Match 75.6%; Score 31; DB 22; Length 302;
Best Local Similarity 85.7%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YTAGOQL 7
|||
Db 79 ylpqgql 85

RESULT 13

AAR66213
ID AAR66213 standard; Protein; 342 AA.

AC AAR66213;

DT 04-AUG-1995 (first entry)

DE Nocardia corallina alkene monooxygenase subunit-1.

XX alkene monooxygenase; subunit-1; epoxidation; amoA gene.

XX Nocardia corallina.

PN JP06292571-A.

PD 21-OCT-1994.

PP 06-APR-1993; 93JP-0105171.

PR 06-APR-1993; 93JP-0105171.

PA (NIHA) JAPAN ENERGY CORP.

DR WPI; 1995-009069/02.

XX N-PSDB; AAQ79569.

PT Alkene monooxygenase and corresp. gene - useful for the
PT epoxidation of an alkene

PS Claim 1; Page 2; 30pp; Japanese.

CC E.coli transformed with the DNA sequence AAQ79569 are able to
CC catalyse the epoxidation of alkenes. The DNA is derived from
CC Nocardia corallina and comprises 4 open reading frames. ORFs amoA
CC and amoC encode subunits 1 and 2 of the alkene monooxygenase enzyme
CC and ORF amod encodes a reductase capable of transferring electrons
CC from NADH coenzyme to a monooxygenase.

Sequence 342 AA;

Query Match 75.6%; Score 31; DB 16; Length 342;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YTAGOQLT 8
|||
Db 34 ylvqgqst 41

RESULT 14

AAR81469
ID AAR81469 standard; Protein; 343 AA.

AC AAR81469;

DT 07-AUG-1996 (first entry)

DE Nocardia corallina alkene mono-oxygenase gene product, amoA.

XX Alkene mono-oxygenase; indole; indigo production; biosynthesis;
XX microbial oxidation; dye.

OS Nocardia corallina B-276.

PN JP08023988-A.

XX 30-JAN-1996.
PD
XX
XX 08-JUL-1994; 94JP-0179688.
PF
XX
XX 08-JUL-1994; 94JP-0179688.
PR
XX
XX (NTHA) JAPAN ENERGY CORP.
PA
XX WPI; 1996-133426/14.
DR
XX N-PSDB; AAT17418.
DR
XX
XX Prep. of indigo by a microbiological method - by culturing a
PT microbe having alkene monooxygenase activity to oxidise indole to
PT indigo
PS
XX Claim 5; Page 6-7; 11pp; Japanese.
XX
XX AAR81469-R81472 are protein products of the Nocardia corallina strain
CC B-276 alkene mono-oxygenase gene. The gene encodes 4 protein products
CC amoa, amob, amoc and amod derived from the 3 different reading frames
CC of the operon. The gene is useful for the production of indigo via
CC oxidation of indole. Nocardia corallina can be cultured in a medium
CC contg. indole and will readily oxidise the indole yielding indigo into
CC the culture medium. E. coli may also be transformed with the alkene
CC mono-oxygenase gene and used as above to efficiently produce indigo
CC by microbial oxidation.
XX
XX
SQ Sequence 343 AA;

Query Match 75.6%; Score 31; DB 17; Length 343;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 YTAGOQLT 8
||| |
Db 35 ylvqqst 42

RESULT 15
AAG64423
ID AAG64423 standard; Protein; 746 AA.
XX
AC AAG64423;
XX
DT 19-SEP-2001 (first entry)
XX
XX Chimeric CBD-fused GL-7ACA acylase 1.
DE
XX
XX Heteromer peptide; chitin cellulose binding domain; CBD;
KW protein immobilisation.
XX
XX Chimeric - Bacillus circulans.
OS
OS Chimeric - unidentified.
XX
XX WO200131038-A1.
PN
XX
PD 03-MAY-2001.
XX
XX 19-OCT-2000; 2000WO-JP07275.
PF
XX
XX 22-OCT-1999; 99JP-0301699.
PR
XX
XX (FUJI) FUJISAMA PHARM CO LTD.
PA
XX
XX Tanaka A, Ueda M, Nagao K;
PI
XX
XX WPI; 2001-417561/44.
DR
XX N-PSDB; AAH46095.
XX
XX Heteromer peptides produced by auto-splicing of a precursor comprise
PT subunits with chitin cellulose binding domain -

XX Claim 8; Page 68-71; 100pp; Japanese.
PS
XX
XX The present sequence is that of the CBD-fused GL-7ACA acylase chimeric
CC protein of the invention. The invention relates to heteromer peptides;
CC produced by cleaving precursor peptides, comprise at least one subunit
CC with an added chitin cellulose binding domain (CBD) useful for
CC immobilising proteins for industrial use. The CBD fusion allows
CC immobilisation of proteins without loss of activity and immobilised
CC heteromer peptides may be produced with high yield.
XX
XX
SQ Sequence 746 AA;

Query Match 75.6%; Score 31; DB 22; Length 746;
Best Local Similarity 75.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 YTAGOQLT 8
||| | : |
Db 16 ytagqlvt 23

Search completed: June 13, 2002, 08:53:30
Job time: 280 sec

Fri Jun 14 10:36:48 2002

us-09-710-339-2_copy_448_455.rag

Page 9

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:54:29 ; Search time 57.95 Seconds
(without alignments)
5.479 Million cell updates/sec

Title: US-09-710-339-2_COPY_98_110
Perfect score: 78
Sequence: 1 AYHGWOODIYSL 13

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCUTS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	78	100.0	478 1	US-08-720-899-7
2	78	100.0	478 1	US-08-459-610-7
3	78	100.0	478 2	US-08-343-804-7
4	78	100.0	478 2	US-08-600-808A-10
5	78	100.0	478 3	US-08-683-838A-10
6	78	100.0	478 4	US-08-182-859-7
7	62	79.5	478 2	US-08-339-715A-2
8	54	69.2	468 1	US-08-204-656B-2
9	54	69.2	468 1	US-08-470-702-6
10	54	69.2	468 1	US-08-467-831-6
11	53	67.9	468 1	US-08-204-656B-4
12	53	67.9	468 1	US-08-470-702-7
13	53	67.9	468 1	US-08-467-831-7
14	50	64.1	468 1	US-08-204-656B-6
15	50	64.1	468 1	US-08-470-702-8
16	49	62.8	468 1	US-08-467-831-8
17	49	62.8	468 1	US-08-204-656B-8
18	49	62.8	468 1	US-08-470-702-9
19	49	62.8	468 1	US-08-467-831-9
20	49	62.8	685 3	US-08-947-965-72
21	49	62.8	685 3	US-08-947-965-74
22	49	62.8	686 3	US-08-947-965-70
23	49	62.8	686 3	US-08-947-965-73
24	47	60.3	624 3	US-08-947-965-78
25	47	60.3	655 1	US-08-469-202-27
26	47	60.3	655 1	US-08-469-202-28
27	47	60.3	655 2	US-08-484-434C-34

28	47	60.3	655 2	US-08-484-434C-35	Sequence 35, Appl
29	46	59.0	675 3	US-08-947-965-76	Sequence 76, Appl
30	46	59.0	676 3	US-08-947-965-71	Sequence 71, Appl
31	46	59.0	680 3	US-08-947-965-77	Sequence 77, Appl
32	46	59.0	683 3	US-08-947-965-72	Sequence 2, Appl
33	46	59.0	687 3	US-08-947-965-75	Sequence 75, Appl
34	46	59.0	725 2	US-08-816-105A-1	Sequence 1, Appl
35	45	57.7	719 4	US-09-386-607-2	Sequence 11, Appl
36	44	56.4	14 1	US-08-204-656B-11	Sequence 11, Appl
37	44	56.4	14 1	US-08-470-702-11	Sequence 11, Appl
38	44	56.4	14 1	US-08-467-831-11	Sequence 11, Appl
39	44	56.4	454 1	US-07-930-686-10	Sequence 10, Appl
40	44	56.4	454 2	US-08-460-998-10	Sequence 10, Appl
41	44	56.4	833 4	US-09-514-302-3	Sequence 3, Appl
42	44	56.4	1938 4	US-09-514-302-2	Sequence 2, Appl
43	43	55.1	14 1	US-08-204-656B-12	Sequence 12, Appl
44	43	55.1	14 1	US-08-470-702-12	Sequence 12, Appl
45	43	55.1	14 1	US-08-467-831-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-720-899-7
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchart, Toden Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowrey Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054,214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-720-899-7

Query Match 100.0%; Score 78; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWOODIYSL 13
 |||
 DB 78 AYHGYWOODIYSL 90

RESULT 2
 US-08-459-610-7
 ; Sequence 7, Application US/08459610
 ; Patent No. 5801043
 ; GENERAL INFORMATION:
 ; APPLICANT: Bisgaard-Frantzen, Henrik
 ; APPLICANT: Borcherdt, Torben Vedel
 ; APPLICANT: Svendsen, Allan
 ; APPLICANT: Thelleresen, Marianne
 ; APPLICANT: Van der Zee, Pia
 ; TITLE OF INVENTION: AMYLASE VARIANTS
 ; NUMBER OF SEQUENCES: 38
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 58010430 No. 5801043disk of No. 5801043th America, Inc.
 ; STREET: 405 Lexington Avenue, 64th Floor
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentln Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/459,610
 ; FILING DATE: 02-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/343,804
 ; FILING DATE: 22-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney Dr., Karen A.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 4054.214-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 478 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-459-610-7

Query Match 100.0%; Score 78; DB 1; Length 478;
 Best Local Similarity 100.0%; Pred. No. 5.2e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWOODIYSL 13
 |||
 DB 78 AYHGYWOODIYSL 90

RESULT 3
 US-08-343-804-7
 ; Sequence 7, Application US/08343804
 ; Patent No. 5830837
 ; GENERAL INFORMATION:
 ; APPLICANT: Bisgaard-Frantzen, Henrik
 ; APPLICANT: Borcherdt, Torben Vedel
 ; APPLICANT: Svendsen, Allan
 ; APPLICANT: Thelleresen, Marianne
 ; APPLICANT: Van der Zee, Pia
 ; TITLE OF INVENTION: AMYLASE VARIANTS

NUMBER OF SEQUENCES: 38
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 58308370 No. 5830837disk of No. 5830837th America, Inc.
 ; STREET: 405 Lexington Avenue, 64th Floor
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentln Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/343,804
 ; FILING DATE: 22-NOV-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney Dr., Karen A.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 4054.214-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 478 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-343-804-7

Query Match 100.0%; Score 78; DB 2; Length 478;
 Best Local Similarity 100.0%; Pred. No. 5.2e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGYWOODIYSL 13
 |||
 DB 78 AYHGYWOODIYSL 90

RESULT 4
 US-08-600-908A-10
 ; Sequence 10, Application US/08600908A
 ; Patent No. 5989169
 ; GENERAL INFORMATION:
 ; APPLICANT: Svendsen, Allan
 ; APPLICANT: Bisg rd-Frantzen, Henrik
 ; APPLICANT: Borcherdt, Torben Vedel
 ; TITLE OF INVENTION: '-Amylase Mutants
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 59891690 No. 5989169disk of No. 5989169th America, Inc.
 ; STREET: 405 Lexington Avenue, 64th Floor
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentln Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/600,908A
 ; FILING DATE: 13-FEB-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Green, Reza
 ; REGISTRATION NUMBER: 38,475
 ; REFERENCE/DOCKET NUMBER: 4394.204-US
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-600-908A-10

Query Match 100.0%; Score 78; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
|||||
DB 78 AYHGWOODIYSL 90

RESULT 5
US-08-683-838A-10
; Sequence 10, Application US/08683838A
; Patent No. 6022724
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: "Amylase Mutants"
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 60227240 No. 6022724disk of No. 6022724th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,838A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/600,908
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394,204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-683-838A-10

Query Match 100.0%; Score 78; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AYHGWOODIYSL 13
|||||

DB 78 AYHGWOODIYSL 90

RESULT 6
US-09-182-859-7
; Sequence 7, Application US/09182859
; Patent No. 6143708
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisgaard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796,204-US
; CURRENT APPLICATION NUMBER: US/09/182,859
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 0515/96
; EARLIER FILING DATE: 1996-04-30
; EARLIER APPLICATION NUMBER: 0712/96
; EARLIER FILING DATE: 1996-06-28
; EARLIER APPLICATION NUMBER: 0775/96
; EARLIER FILING DATE: 1996-07-11
; EARLIER APPLICATION NUMBER: 1263/96
; EARLIER FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-182-859-7

Query Match 100.0%; Score 78; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
|||||
DB 78 AYHGWOODIYSL 90

RESULT 7
US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michio
; APPLICANT: Takata, Hiroki
; APPLICANT: Okada, Shigetaka
; TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
; NUMBER OF SEQUENCES: 7
; TITLE OF INVENTION: NEOPULULANASE
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POHMAN, HAIK, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; STREET: G. Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,715A
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Melsner, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335.009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 79.5%; Score 62; DB 2; Length 478;
Best Local Similarity 84.6%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGYWOODIYSL 13
|||:|:|:|:|:|
Db 78 ATGTGWQDIYSL 90

RESULT 8
US-08-204-656B-2
Sequence 2, Application US/08204656B
Patent No. 5538882
GENERAL INFORMATION:
APPLICANT: Matsui, Ikuro
APPLICANT: Ishikawa, Kazuhiko
APPLICANT: Miyairi, Sachio
APPLICANT: Honda, Koichi
TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
TITLE OF INVENTION: Oligosaccharide Using The Enzyme
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-204-656B-2

Query Match 69.2%; Score 54; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.34;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGYWOODIYSL 13
|||:|:|:|:|:|
Db 79 AYHGFWMKNIYKI 91

RESULT 9
US-08-470-702-6
Sequence 6, Application US/08470702
Patent No. 5631149
GENERAL INFORMATION:
APPLICANT: MATSUI, IKURO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,702
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-470-702-6

Query Match 69.2%; Score 54; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.34;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGYWOODIYSL 13
|||:|:|:|:|:|
Db 79 AYHGFWMKNIYKI 91

RESULT 10
US-08-467-831-6
; Sequence 6, Application US/08467831
; Patent No. 5635378
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,831
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,656
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-467-831-6

Query Match 69.2%; Score 54; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.34;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL.13
||||:|:|:|:
Db 79 AYHGFMKNITIKI 91

RESULT 11
US-08-204-656B-4
; Sequence 4, Application US/08204656B
; Patent No. 5518882
; GENERAL INFORMATION:
; APPLICANT: Matsui, Ikuo
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Miyairi, Sachio
; APPLICANT: Honda, Koichi
; TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
; TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
; TITLE OF INVENTION: Oligosaccharide Using The Enzyme

; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,656B
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-204-656B-4

Query Match 67.9%; Score 53; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.49;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
||||:|:|:|:
Db 79 AYHGFMKNITIKI 91

RESULT 12
US-08-470-702-7
; Sequence 7, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,702
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-470-702-7

Query Match 67.9%; Score 53; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.49;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGWOODIYSL 13
|||||:|:|:|:
Db 79 AYHGMMKNITYKI 91

RESULT 13
US-08-467-831-7
Sequence 7, Application US/08467831
Patent No. 5635378
GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
NUMBER OF INVENTION: 17
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,831
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:

LENGTH: 468 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-467-831-7

Query Match 67.9%; Score 53; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 0.49;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AYHGWOODIYSL 13
|||||:|:|:|:
Db 79 AYHGMMKNITYKI 91

RESULT 14
US-08-204-656B-6
Sequence 6, Application US/08204656B
Patent No. 5538882
GENERAL INFORMATION:
APPLICANT: Matsui, Ikuo
APPLICANT: Ishikawa, Kazuhiko
APPLICANT: Miyairi, Sachio
APPLICANT: Honda, Koichi
TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
NUMBER OF INVENTION: 18
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/204,656B
FILING DATE: 02-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiner, Marc S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-204-656B-6

Query Match 64.1%; Score 50; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 1.5;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AYHGWOODIYSL 13
|||||:|:|:|:
Db 79 AYHGMMKNITYKI 91

RESULT 15
US-08-470-702-8
; Sequence 8, Application US/08470702
; Patent No. 5631149
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: MIYAIRI, SACHIO
; APPLICANT: HONDA, KOICHI
; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 810 Gatehouse Road, Suite 500 East
; City: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,702
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/204,656
; FILING DATE: 02-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 234-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-470-702-8

Query Match 64.1%; Score 50; DB 1; Length 468;
Best Local Similarity 53.8%; Pred. No. 1.5;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AYHGTYQDDIYSL 13
|||||::|||:
Db 79 AYHGLMKNIYKI 91

Search completed: June 13, 2002, 08:54:29
Job time: 314 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:06 ; Search time 125.81 Seconds
(without alignments)

17.876 Million cell updates/sec

Title: US-09-710-339-2_COPY_98_110

Sequence: 1 AHHGYWQDIYSL 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	78	100.0	499	3	Q96TH4	Q96TH4 aspergillus
2	75	96.2	490	3	Q9UV07	Q9UV07 emericella
3	68	87.2	128	5	015751	015751 dictyostell
4	66	84.6	631	3	Q92394	Q92394 cryptococcu
5	64	82.1	507	3	Q08806	Q08806 debaryomyce
6	63	80.8	640	3	013296	013296 aspergillus
7	62	79.5	624	3	001117	001117 lipomyces k
8	59	75.6	482	2	060051	060051 thermosactin
9	54	69.2	491	3	013396	013396 schizosacch
10	54	69.2	623	3	Q9UV09	Q9UV09 emericella
11	53	67.9	564	3	Q9Y7S9	Q9Y7S9 schizosacch
12	52	66.7	711	16	Q99ZB3	Q99ZB3 streptococc
13	50	64.1	625	3	074922	074922 schizosacch
14	49	62.8	690	16	Q9KLB6	Q9KLB6 vibrio chol
15	49	62.8	713	2	Q9FSW3	Q9FSW3 bacillus ci
16	47	60.3	483	16	Q9RUB8	Q9RUB8 deinococcus

17	46	59.0	692	2	Q30565	Q30565 bacillus br
18	46	59.0	704	2	Q82984	Q82984 bacillus sp
19	46	59.0	711	2	Q9ZAQ0	Q9ZAQ0 bacillus st
20	46	59.0	714	2	Q52766	Q52766 penibacill
21	46	59.0	724	2	Q9K515	Q9K515 actinoplan
22	46	59.0	725	2	Q59239	Q59239 bacillus sp
23	45	57.7	739	1	Q9UWN2	Q9UWN2 thermococcu
24	44	56.4	459	5	Q45132	Q45132 haemophilus
25	44	56.4	677	2	Q9RHR1	Q9RHR1 klebsiella
26	44	56.4	1938	2	P70983	P70983 bacillus sp
27	43	55.1	134	5	Q9MSU9	Q9MSU9 drosophila
28	43	55.1	158	5	Q965N4	Q965N4 caenorhabdi
29	43	55.1	464	4	Q9NSY6	Q9NSY6 homo sapien
30	43	55.1	476	11	Q99LV4	Q99LV4 mus musculu
31	43	55.1	552	10	Q9LVM8	Q9LVM8 arabidopsis
32	43	55.1	612	4	Q96K36	Q96K36 homo sapien
33	43	55.1	768	4	Q96K54	Q96K54 homo sapien
34	43	55.1	793	4	Q96K44	Q96K44 homo sapien
35	43	55.1	793	4	Q96K38	Q96K38 homo sapien
36	43	55.1	793	11	Q9DC23	Q9DC23 mus musculu
37	42	53.8	198	5	Q9VRM1	Q9VRM1 drosophila
38	42	53.8	275	2	Q9ALM7	Q9ALM7 saccharopol
39	42	53.8	397	16	Q913T2	Q913T2 pseudomonas
40	41	52.6	153	2	Q93A49	Q93A49 enterococcu
41	41	52.6	410	13	Q9PW33	Q9PW33 typhloocte
42	41	52.6	413	10	Q9M6R9	Q9M6R9 malus domes
43	41	52.6	517	2	Q9XRB6	Q9XRB6 streptomyce
44	41	52.6	1021	2	Q9K5L6	Q9K5L6 actinoplan
45	41	52.6	1798	2	Q9K211	Q9K211 streptomyce

ALIGNMENTS

RESULT	ID	Q96TH4	PRELIMINARY	PRT	499 AA
AC	Q96TH4				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	TAKA-AMYLASE A (EC 3.2.1.1) (ALPHA-AMYLASE).				
GN	AMYA OR AMY1.				
OS	Aspergillus oryzae, and				
OC	Aspergillus flavus.				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				
OX	NCBI_TaxID=5062, 5059;				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=A.oryzae; STRAIN=R1840;				
RX	MEDLINE=20289310; PubMed=10830498;				
RA	Gomi K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N.,				
RA	Iimura Y.;				
RT	"Molecular cloning and characterization of a transcriptional activator				
RT	gene, amyR, involved in the amylolytic gene expression in Aspergillus				
RT	oryzae."				
RL	Biosci. Biotechnol. Biochem. 64:816-827(2000).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=A.flavus; STRAIN=86-10D;				
RA	Fakhoury A.M., Woloshuk C.P.;				
RT	"Amyl, the alpha-amylase gene of Aspergillus flavus: Involvement in				
RT	afatoxin biosynthesis in maize kernels."				
RL	Phytopathology 89:908-914(1999).				
DR	EMBL; AB021876; BA85703.1; -.				
DR	EMBL; AF139925; AF14264.1; -.				
SO	SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;				
Query Match	100.0%;	Score 78;	DB 3;	Length 499;	
Best Local Similarity	100.0%;	Pred. No. 0.00015;			
Matches 13;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	

OY 1 AYHGWOODIYSL 13
|||||
DB 99 AYHGWOODIYSL 111

RESULT 2
ID Q9UV07 PRELIMINARY; PRT; 490 AA.
AC Q9UV07;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE AMYA.
GN AMYA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RA Boase N.A., Murphy R.M., Kelly J.M.;
RT "An amylase cluster in Aspergillus nidulans."
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF208225; AAF17103.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02806; alpha-amylase_C; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
SQ SEQUENCE 490 AA; 54249 MW; A891C4ACEABE5305 CRC64;

Query Match 96.2%; Score 75; DB 3; Length 490;
Best Local Similarity 92.3%; Pred. No. 0.00043;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
|||||
DB 91 AYHGWOODIYSL 103

RESULT 3
ID 015751 PRELIMINARY; PRT; 128 AA.
AC 015751;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE AMYA (FRAGMENT).
GN AMYA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Loomis W.F., Iraifar N.;
RT Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF020284; AAB70852.1; -.
DR HSSP; P10529; 7TAA.
FT NON_TER 1
FT NON_TER 128
SQ SEQUENCE 128 AA; 14950 MW; 2ACB8FDC55E79637 CRC64;

Query Match 87.2%; Score 68; DB 5; Length 128;
Best Local Similarity 83.3%; Pred. No. 0.0014;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 YHGWOODIYSL 13
|||||
DB 3 YHGWOODIYTV 14

RESULT 4
ID Q92394 PRELIMINARY; PRT; 631 AA.
AC Q92394;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ALPHA-AMYLASE PRECURSOR.
GN AMY-CS2.
OS Cryptococcus sp. S-2.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
OC mitosporic Hymenomycetes; Cryptococcus.
OX NCBI_TaxID=87049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S-2;
RX MEDLINE=96433120; PubMed=8836148;
RA Refuji H., Chino M., Kato M., Jimura Y.;
RT "Raw starch digesting and thermostable alpha-amylase from the yeast
RT Cryptococcus sp. S-2: purification, characterization, cloning and
RT sequencing.";
RT Biochem. J. 318:989-996(1996).
DR EMBL; D83541; BAA12011.1; -.
DR EMBL; D83540; BAA12010.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR002044; CBD_4.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF00686; CBD_4; 1.
DR Prodom; PD001568; CBD_4; 1.
KW Signal.
FT SIGNAL 1 20
FT CHAIN 21 631
FT SEQUENCE 631 AA; 67658 MW; 8196B7B6E1D707E5 CRC64;

Query Match 84.6%; Score 66; DB 3; Length 631;
Best Local Similarity 76.9%; Pred. No. 0.016;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
|||||
DB 105 AYHGWOODIYET 117

RESULT 5
ID Q08806 PRELIMINARY; PRT; 507 AA.
AC Q08806;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ALPHA-AMYLASE 2 (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE).
GN SMA2.
OS Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=27300;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 26077;
RX MEDLINE=93365041; PubMed=8358835;
RA Claretos M.G., Abatca D., Fernandez-Lobato M., Jimenez A.;
RT "Molecular structure of the SMA2 gene encoding an AMY1-related alpha-
RT amylase from Schwanniomyces occidentalis.";
RT Curr. Genet. 24:75-83(1993).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- ENZYME REGULATION: ALPHA-AMYLASE EXPRESSION UNDERLIES CARABOLITE
CC REPRESSION BY GLUCOSE.

CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 DR EMBL: X73497; CAA51912.1; -.
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha_amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KM Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KM Glycoprotein.
 FT ACT_SITE 238 238 BY SIMILARITY.
 FT ACT_SITE 242 242 BY SIMILARITY.
 FT ACT_SITE 329 329 BY SIMILARITY.
 FT DISULFID 62 70 BY SIMILARITY.
 FT DISULFID 182 196 BY SIMILARITY.
 FT DISULFID 272 315 BY SIMILARITY.
 FT DISULFID 470 505 BY SIMILARITY.
 FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (PROBABLE).
 SQ SEQUENCE 507 AA; 55966 MW; 3A562E95BD8AD63 CRC64;
 Query Match 82.1%; Score 64; DB 3; Length 507;
 Best Local Similarity 69.2%; Pred. No. 0.026; 1; Indels 0; Gaps 0;
 Matches 9; Conservative 3; Mismatches 1;
 OY 1 AYHGWMKDLYSL 13
 DB 110 AYHGWMKDLYSL 122
 RESULT 6
 ID 013296 PRELIMINARY; PRT; 640 AA.
 AC 013296;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ACID-STABLE ALPHA-AMYLASE.
 OS Aspergillus kawachii.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eukaryotes; Trichocomaceae; mitosporic Trichocomaceae; Aspergillius.
 OX NCBI_TaxID=40384;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kaneko A.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RA Kaneko A.; Sudo S.; Sakamoto Y.; Tamura G.; Ishikawa T.; Ohba T.;
 RT "Molecular-cloning and determination of the nucleotide-sequence of a
 RT gene encoding an acid-stable alpha-amylase from Aspergillus-
 RT kawachii.";
 RL U. Ferment. Bioeng. 81:292-298(1996).
 DR EMBL: AB008370; BA022993.1; -.
 DR HSSP: P56271; ZAAA.
 DR InterPro: IPR000461; Alpha_amylase.
 DR InterPro: IPR002044; CBD_4.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR ProDom: PD001568; CBD_4; 1.
 DR PROSITE: PS00012; alpha-amylase; 1.
 SQ SEQUENCE 640 AA; 69507 MW; 676BF9D0236720DD CRC64;
 Query Match 80.8%; Score 63; DB 3; Length 640;
 Best Local Similarity 90.9%; Pred. No. 0.048;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 AYHGWMKDLYSL 11
 DB 99 AYHGWMKDLYSL 109
 RESULT 7
 ID 001117 PRELIMINARY; PRT; 624 AA.
 ID 001117 PRELIMINARY; PRT; 624 AA.

AC 001117;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1)
 DE (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE) (UKAI).
 GN UKAI.
 OS Lipomyces kononenkoae.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Lipomyetaceae; Lipomyces.
 OX NCBI_TaxID=34357;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IGC4052B;
 RA MEDLINE=96105202; PubMed=8529895;
 RX Steyn A.J.C.; Marmur J.; Pretorius I.S.;
 RT "Cloning, sequence analysis and expression in yeasts of a cDNA
 RT containing a Lipomyces kononenkoae alpha-amylase-encoding gene.";
 RL Gene 166:65-71(1995).
 RN [2]
 RP SEQUENCE OF 29-44.
 RC STRAIN=IGC4052B;
 RX MEDLINE=96132108; PubMed=8593683;
 RA Steyn A.J.C.; Pretorius I.S.;
 RT "Characterization of a novel alpha-amylase from Lipomyces kononenkoae
 RT and expression of its gene (UKAI) in Saccharomyces cerevisiae.";
 RL Curr. Gene. 28:526-533(1995).
 CC -1- FUNCTION: LIBERATES REDUCING GROUPS FROM POLYMERS CONTAINING BOTH
 CC 1,4-ALPHA AND 1,6-ALPHA BONDS.
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 DR EMBL: U30376; AAC49622.1; ALT_INIT.
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha_amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KM Hydrolase; Glycosidase; Carbohydrate metabolism; Signal; Calcium;
 KM Glycoprotein.
 FT SIGNAL 1 28
 FT CHAIN 29 624 ALPHA-AMYLASE.
 FT ACT_SITE 353 353 BY SIMILARITY.
 FT ACT_SITE 357 357 BY SIMILARITY.
 FT ACT_SITE 377 377 BY SIMILARITY.
 FT ACT_SITE 444 444 BY SIMILARITY.
 FT DISULFID 177 185 BY SIMILARITY.
 FT DISULFID 297 311 BY SIMILARITY.
 FT DISULFID 387 430 BY SIMILARITY.
 FT DISULFID 587 622 BY SIMILARITY.
 FT DISULFID 587 622 BY SIMILARITY.
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 624 AA; 68876 MW; 87EB16534F5A99F CRC64;
 Query Match 79.5%; Score 62; DB 3; Length 624;
 Best Local Similarity 69.2%; Pred. No. 0.067;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 1 AYHGWMKDLYSL 13
 DB 225 AYHGWMKDLYSL 237
 RESULT 8
 ID 060051 PRELIMINARY; PRT; 482 AA.
 AC 060051;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1).
 GN AMYV.
 OS Thermocotiniomyces vulgaris.
 OC Bacteria: Firmicutes; Bacillus/Clostridium group; Thermocotiniomyces.
 OX NCBI_TaxID=2026;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=K94;
 RX MEDLINE=95031040; PubMed=7944369;
 RA Hofmeister B., Koenig S., Hoang V., Engel J., Mayer G., Hansen G.,
 Hofmeister J.;
 RT "The gene amy(TY1) codes for a nonfluorogenic alpha-amylase from
 KT Thermocotiniomyces vulgaris 94-2A in Bacillus subtilis."
 RL Appl. Environ. Microbiol. 60:3381-3389(1994).
 DR EMBL; X69807; CAA49465.1; -.
 DR HSSP; P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02806; alpha-amylase-C; 1.
 KW Signal; Hydrolase; Glycosidase.
 FT SIGNAL 1 29 POTENTIAL.
 FT CHAIN 30 482 ALPHA-AMYLASE.
 FT SEQUENCE 482 AA; 54483 MW; D8F98C370B579025 CRC64;

Query Match 75.6%; Score 59; DB 2; Length 482;
 Best Local Similarity 69.2%; Pred. No. 0.16;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 AYHGYYQODIYSL 13
 Db 107 AYHGYYQDYFYSV 119
 ||||| :|||

RESULT 9
 ID 013996 PRELIMINARY; PRT; 491 AA.
 AC 013996;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PUTATIVE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1).
 GN SPAC27E2.O1.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Wood V.,
 RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 DR EMBL; Z69878; CAB11675.1; -.
 DR HSSP; P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
 KW Glycoprotein; Signal.
 FT SIGNAL 1 29 POTENTIAL.
 FT CHAIN 30 491 PUTATIVE ALPHA-AMYLASE.
 FT ACT_SITE 221 221 BY SIMILARITY.
 FT ACT_SITE 225 225 BY SIMILARITY.
 FT ACT_SITE 313 313 BY SIMILARITY.
 FT CARBOHYD 307 307 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 319 319 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 391 391 N-LINKED (GLCNAc. . .) (POTENTIAL).
 SO SEQUENCE 491 AA; 56678 MW; 2013808DC0EE581D CRC64;

Query Match 69.2%; Score 54; DB 3; Length 491;
 Best Local Similarity 61.5%; Pred. No. 1;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 AYHGYYQODIYSL 13
 Db 94 AYHGYYQADLTOL 106
 ||||| :|||

RESULT 10
 ID 09YV09 PRELIMINARY; PRT; 623 AA.
 AC 09YV09;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ALPHA-AMYLASE.
 GN AMYB.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=5072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Boase N.A., Murphy R.L., Kelly J.M.;
 RT "Characterisation and regulation of an alpha-amylase gene in
 RT Aspergillus nidulans.";
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF208224; AAF17100.1; -.
 DR HSSP; P56271; 2AAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR InterPro: IPR002044; CBD_4.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR PRINTS; PR00110; ALPHAAMYLASE.
 DR PRODOM; PD001566; CBD_4; 1.
 SO SEQUENCE 623 AA; 68598 MW; 822F002C37F5A9A4 CRC64;

Query Match 69.2%; Score 54; DB 3; Length 623;
 Best Local Similarity 58.3%; Pred. No. 1.3;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 YHGYYQODIYSL 13
 Db 98 YHGYYQKNYGV 109
 :||| :|||

RESULT 11
 ID 09Y7S9 PRELIMINARY; PRT; 564 AA.
 AC 09Y7S9;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PROBABLE ALPHA-AMYLASE C63.02C PRECURSOR (EC 3.2.1.1) (1,4-ALPHA-D-
 DE GLUCAN GLUCANOHYDROLASE).
 GN SPC63.02C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Murphy L., Harris D., Wood V., Rajandream M.A., Barrell B.G.,
 RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
 CC SIMILARITY).

CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 DR EMBL: AL049522; CAB40006.1; -
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Hypothetical protein; Hydrolase; glycosidase; Carbohydrate metabolism;
 KW Calcium; Glycoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 564
 FT DOMAIN 537 540
 FT ACT_SITE 229 229
 FT ACT_SITE 322 322
 FT DISULFID 51 59
 FT DISULFID 172 188
 FT DISULFID 263 306
 FT CARBOHYD 181 181
 FT CARBOHYD 235 235
 FT CARBOHYD 282 282
 FT CARBOHYD 305 305
 FT CARBOHYD 438 438
 FT CARBOHYD 447 447
 FT CARBOHYD 498 498
 SQ SEQUENCE 564 AA; 63205 MW; 1229FD4AEC702FED0 CRC64;

Query Match 67.9%; Score 53; DB 3; Length 564;
 Best Local Similarity 69.2%; Pred. No. 1.7;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
 ||||| 1 11
 Db 100 AYHGWMNTDESL 112

RESULT 12
 Q99ZB3 PRELIMINARY; PRT; 711 AA.
 AC Q99ZB3;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE CYCLOMALTODEXTRIN GLUCANOTRANSFERASE (EC 2.4.1.19).
 GN AMYA OR SPY1302.
 OS Streptococcus pyogenes.
 OC Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_Taxid=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;
 RX MEDLINE=21192684; PubMed=11296296;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 DR EMBL: AE006569; AAK3149.1; -
 DR HSSP: P4379; 1CDG.
 DR InterPro: IPR00461; Alpha-amylase.
 DR InterPro: IPR002453; Beta-tubulin.
 DR InterPro: IPR002044; CBD_4.
 DR InterPro: IPR002909; IPT_TIG.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02806; alpha-amylase; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR Pfam: PF01833; TIG; 1.
 DR PRINTS: PR00110; ALPHAAMYLAZE.
 DR PRODOM: PD001568; CBD_4; 1.
 DR PROSITE: PS00228; TUBULIN_B-ATOREG; UNKNOWN_1.
 DR Transference; Glycosyltransferase; Complete proteome.
 KW TRANSFERENCE; 711 AA; 79544 MW; 88F7AA9424C0E717 CRC64;
 SQ SEQUENCE

Query Match 66.7%; Score 52; DB 16; Length 711;
 Best Local Similarity 63.6%; Pred. No. 3.1;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYHGWOODIY 11
 ||||| 1 1;
 Db 122 AYHGWMNDFF 132

RESULT 13
 Q74922 PRELIMINARY; PRT; 625 AA.
 AC Q74922;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ALPHA-AMYLASE.
 GN SPC0757.12.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_Taxid=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972H-;
 RA Lyne M., Rajadream M.A., Barrell B.G., Bothe G., Pohl T;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL031825; CAA21237.1; -
 DR HSSP: P10529; 7TAA
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 SQ SEQUENCE 625 AA; 68087 MW; 30A58CE6D6BC07B3 CRC64;

Query Match 64.1%; Score 50; DB 3; Length 625;
 Best Local Similarity 61.5%; Pred. No. 5.7;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
 ||||| 1 1
 Db 100 AYHGWMNTDFQL 112

RESULT 14
 Q9KLB6 PRELIMINARY; PRT; 690 AA.
 ID Q9KLB6
 AC Q9KLB6;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ALPHA-AMYLASE.
 GN VCA0860.
 OS Vibrio cholerae.
 OC Bacteria: Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
 OX NCBI_Taxid=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.D., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae.";
 RL Nature 406:477-483(2000).
 DR EMBL: AE004413; AAF96758.1; -

DR TIGR: VCA0860; -
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Complete proteome.
 SQ SEQUENCE 690 AA; 77331 MW; 7C82C00C928FDE41 CRC64;

Query Match 62.8%; Score 49; DB 16; Length 690;
 Best Local Similarity 77.8%; Pred. No. 9.1;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGYWQD 9
 ||||| :|
 Db 288 AYHGYWTRD 296

RESULT 15

O9F5W3 PRELIMINARY; PRT; 713 AA.
 AC O9F5W3.
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CYCLODEXTRIN GLUCANOSYLTRANSFERASE.
 OS Bacillus circulans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1397;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-All;
 RA Rhiphanitchayakit V., Tonozuka T., Sakano Y.;
 RT "Cloning of Cyclodextrin Glucanoyltransferase Gene from Bacillus
 RT circulans All.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF302787; AAC31622.1; -.
 DR HSSP: P05618; 1PAM.
 DR InterPro: IPR000461; Alpha-amylase.
 DR InterPro: IPR002044; CBD_4
 DR InterPro: IPR002909; IPT_TIG.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02806; alpha-amylase_C; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR Pfam: PF01833; TIG; 1.
 DR PRINTS: PR00110; ALPHAAMYLASE.
 DR ProDom: PD001568; CBD_4; 1.
 KW Transferrase.
 SQ SEQUENCE 713 AA; 78271 MW; 4FF365E1F8F93BAD CRC64;

Query Match 62.8%; Score 49; DB 2; Length 713;
 Best Local Similarity 77.8%; Pred. No. 9.4;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGYWQD 9
 ||||| :|
 Db 123 AYHGYWARD 131

Search completed: June 13, 2002, 09:11:07
 Job time: 1052 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:48 ; Search time 38.1 Seconds

(without alignments)
13.211 Million cell updates/sec

Title: US-09-710-339-2_COPY_98_110
Perfect score: 78
Sequence: 1 AHHGYWQDIYSL 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	78	100.0	498	1 AMYA_ASPAW	Q02905 aspergillus
2	78	100.0	499	1 AMYA_ASPOR	P10529 aspergillus
3	78	100.0	499	1 AMYA_ASPAW	Q02906 aspergillus
4	78	100.0	499	1 AMY_ASPSH	P30282 aspergillus
5	67	85.9	513	1 AMY2_SCHPO	O14154 schizosacch
6	64	82.1	484	1 AMYA_ASPNG	P56271 aspergillus
7	58	74.4	494	1 AMY1_SACFI	P21567 saccharomyc
8	56	71.8	478	1 YDOL_SCHPO	O10427 schizosacch
9	52	66.7	581	1 AMY1_SCHPO	Q09840 schizosacch
10	50	64.1	512	1 AMY1_DEBOC	P19269 debaryomyc
11	49	62.8	528	1 AMY_BACCT	P08137 bacillus ci
12	49	62.8	712	1 CDGT_BACS3	P09121 bacillus sp
13	49	62.8	713	1 AMYR_BACS8	P17692 bacillus sp
14	49	62.8	713	1 CDG2_PAPMA	P31835 paenibacill
15	49	62.8	713	1 CDGT_BACSO	P05618 bacillus sp
16	49	62.8	713	1 CDGT_BACSP	P30921 bacillus sp
17	49	62.8	713	1 CDGT_BACCI	P43379 bacillus ci
18	49	62.8	718	1 CDGT_BACCI	P30920 bacillus ci
19	49	62.8	718	1 CDGT_BACLI	P14014 bacillus li
20	49	62.8	718	1 CDGT_BACSS	P31747 bacillus sp
21	47	60.3	655	1 CDGT_KLEPN	P08704 klebsiella
22	46	59.0	703	1 CDGT_BACS2	P31746 bacillus sp
23	46	59.0	704	1 CDGT_BACOS	P27036 bacillus oh
24	46	59.0	710	1 CDGT_THETU	P26827 thermomanc
25	46	59.0	711	1 CDGT_BACST	P31797 bacillus st
26	46	59.0	714	1 CDG1_PAPMA	P04830 paenibacill
27	45	57.7	717	1 AMYM_BACST	P19531 bacillus st
28	44	56.4	676	1 AMY1_ECOLI	P25718 escherichia
29	44	56.4	2158	1 MY9B_HUMAN	Q13459 homo sapien
30	43	55.1	1196	1 AMYB_PAPPO	P21543 paenibacill
31	41	52.6	499	1 MYFD_RHOCA	P08717 rhodospirillum
32	41	52.6	919	1 AMY_STRLI	P05884 streptomyces
33	40.5	51.9	604	1 RCOL_NEUCR	P78706 neurospora

34	40	51.3	508	1	GSN1_MEDTR	Q92NX5 medicago tr
35	40	51.3	734	1	PSAH_CYACA	Q9L196 cyanidium c
36	40	51.3	1270	1	PEAH_ECOLI	P33546 escherichia
37	39	50.0	395	1	RRP_ECOLI	P16073 newcastle d
38	39	50.0	395	1	RRP_NDVA	P24698 newcastle d
39	39	50.0	395	1	RRP_NDV2	P06427 newcastle d
40	38	48.7	383	1	PEPE_CHICK	P16476 gallus gall
41	38	48.7	387	1	PEPE_CHICK	P27822 oryctolagus
42	38	48.7	413	1	HMT_DROME	P02835 drosophila
43	38	48.7	501	1	C6AH_DROME	Q9V770 drosophila
44	38	48.7	502	1	C6AH_DROME	Q9V771 drosophila
45	38	48.7	559	1	CYSI_THIRO	P52673 thiocapsa r

ALIGNMENTS

```

RESULT 1
ID AMYA_ASPAW STANDARD: PRT: 498 AA.
AC 002905:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase A).
GN AMYA.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VVK143P.
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
RT from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X52755; CAA36966.1; -.
CC HSSP: P10529; TTA.
CC InterPro: IPR000461; Alpha.amylase.
CC Pfam: PF00128; alpha-amylase.1.
KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 21
FT CHAIN 1 498
FT ACT_SITE 227 498
FT ACT_SITE 227 227
FT ACT_SITE 231 231
FT ACT_SITE 251 251
FT ACT_SITE 251 251
FT ACT_SITE 318 318
FT ACT_SITE 318 318
FT DISULFID 51 59
FT DISULFID 171 185
FT DISULFID 261 304
FT DISULFID 461 496
FT CARBOHYD 218
FT SEQUENCE 498 AA; 54880 MM; 7658511BC01ABA01 CRC64;
N-LINKED (GLCNAc... ) (POTENTIAL).

```

Query Match 100.0%; Score 78; DB 1; Length 498;
 Best Local Similarity 100.0%; Pred. No. 1,4e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
 |||||
 DB 99 AYHGWOODIYSL 111

RESULT 2
 AMYA_ASPOR STANDARD; PRT: 499 AA.

ID AMYA_ASPOR
 AC P10529; P11763; Q00250;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
 DE (1,4-alpha-D-glucan glucohydrolase).
 GN AMY1 AND AMY2 AND AMY3.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OC NCBI_TaxID=5062;
 RX STRAIN=DSM 63303;
 RX MEDLINE=89237897; PubMed=2785629;
 RX Wirsel S., Lachmund A., Wildhardt G., Rutkowski E.;
 RT "three alpha-amylase genes of Aspergillus oryzae exhibit identical
 RT intron-exon organization.";
 RL Mol. Microbiol. 3:3-14(1989).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89378767; PubMed=2789162;
 RX Genes M.J., Dove M.J., Seliy V.L.;
 RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
 RT containing eight introns.";
 RL Gene 79:107-117(1989).
 RN [13]
 RP SEQUENCE FROM N.A.
 RX Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshikawa K.;
 RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
 RT of Aspergillus oryzae.";
 RL Agric. Biol. Chem. 53:593-599(1989).
 RN [14]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90128276; PubMed=2612911;
 RX Tsukagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A.,
 RA Ueda S.;
 RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
 RT evidence for multiple related genes.";
 RL Gene 84:319-327(1989).
 RN [15]
 RP SEQUENCE OF 22-499.
 RA Toda H., Kondo K., Narita K.;
 RT "The complete amino acid sequence of Taka-amylase A.";
 RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
 RN [16]
 RP SEQUENCE OF 206-225.
 RX MEDLINE=74001521; PubMed=4733850;
 RX Isemura S., Ikenaka T.;
 RA "The amino acid sequences of glycopeptides obtained from Taka-amylase
 RA A with trypsin and chymotrypsin.";
 RL J. Biochem. 74:1-10(1973).
 RN [17]
 RP SEQUENCE OF 433-499.
 RA Narita K.;
 RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 51:285-290(1975).
 RN [18]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=80227691; PubMed=6156152;

RA Matsunura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
 RA Toda H., Narita K., Kakudo M.;
 RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
 RT 3-A resolution.";
 RL J. Biochem. 87:1555-1558(1980).
 RN [19]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=84212370; PubMed=6609921;
 RA Matsunura Y., Kusunoki M., Harada W., Kakudo M.;
 RT "Structure and possible catalytic residues of Taka-amylase A.";
 RL J. Biochem. 95:697-702(1984).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
 RX MEDLINE=97428212; PubMed=9283074;
 RA Brzozowski A.M., Davies G.J.;
 RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
 RT inhibitor acarbose at 2.0-A resolution.";
 RL Biochemistry 36:10837-10845(1997).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- MISCELLANEOUS: THE SEQUENCE OF AMY1 AND AMY2 IS SHOWN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X12725; CAA31218.1; -;
 DR EMBL: X12726; CAA31219.1; -;
 DR EMBL: X12727; CAA31220.1; -;
 DR EMBL: D00434; BAA00336.1; -;
 DR EMBL: M33218; AAA32708.1; -;
 DR PIR: S04348; ALAS1.
 DR PIR: S04549; ALAS3.
 DR PIR: A10627; A10627.
 DR PIR: JK0201; JK0201.
 DR PIR: JS0240; JS0240.
 DR PIR: JTO466; JTO466.
 DR PDB: 2TAA; 15-OCT-89.
 DR PDB: 6TAA; 31-OCT-93.
 DR PDB: 7TAA; 25-NOV-98.
 DR GLYCOSULEDB: P10529; -;
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase: 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 59
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 FT
 FT VARIANT 56 56
 FT VARIANT 172 172
 FT CONFLICT 93 94
 FT CONFLICT 106 106
 FT CONFLICT 184 184
 FT CONFLICT 195 195
 FT CONFLICT 255 255
 FT CONFLICT 291 291
 FT CONFLICT 345 345
 N-LINKED (GLCNAC. . .).
 /FTID=CAR_000125.
 Q -> R (IN AMY3).
 F -> L (IN AMY3).
 T* -> DC (IN REF. 5).
 Q -> T (IN REF. 5).
 D -> Y (IN REF. 3).
 P -> L (IN REF. 3).
 G -> V (IN REF. 3).
 D -> H (IN REF. 4).
 I -> L (IN REF. 5).

Query Match	100.0%;	Score 78;	DB 1;	Length 499;
Best Local Similarity	100.0%;	Pred. No. 1.4e-05;		
Matches 13; Conservative	0;	Mismatches	0;	Indels 0;
Gaps	0;			
0Y	1	AYHGWMQDDIYSL	13	

Query Match	100.0%;	Score 78;	DB 1;	Length 499;
Best Local Similarity	100.0%;	Pred. No. 1,4e-05;		
Matches 13; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AYHGWMODIYSL	13	
Db	99	AYHGWMODIYSL	111	
RESULT	4			
AMY_ASFSH				

```

ID      AMY_ASPEH      STANDARD:      PRT:      499 AA.
AC      P30292;
DT      01-APR-1993 (Rel. 25, Created)
DT      01-APR-1993 (Rel. 25, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Alpha-amylose precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE      glucanohydrolase).
GN      AMY.
OS      Aspergillus shirousami.
OC      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC      Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX      NCBI_TaxID=5070;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=9232146; PubMed=1368777;
RA      Shibuya I., Tamura G., Ishikawa T., Hara S.;
RT      "Cloning of the alpha-amylose cDNA of Aspergillus shirousami and its
RT      expression in Saccharomyces cerevisiae."
RL      Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC      -1 CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC      linkages in oligosaccharides and polysaccharides.
CC      -1 COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC      -1 SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC      KNOWN AS THE ALPHA-AMYLASE FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL: D10461; BAA01255.1; -.
DR      PIR: J50663; J50663.
DR      HSSP: P10529; 7TAA.
DR      InterPro: IPR000461; Alpha-amylose.
DR      Pfam: PF00128; alpha-amylose; 1.
KW      Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW      Glycoprotein; Signal.
FT      SIGNAL      1
FT      CHAIN      22 499
FT      ACT_SITE    227 227
FT      ACT_SITE    231 231
FT      ACT_SITE    251 251
FT      ACT_SITE    318 318
FT      ACT_SITE    51 59
FT      DISULFID    171 185
FT      DISULFID    261 304
FT      DISULFID    461 496
FT      CARBOHYD    218 218
SQ      SEQUENCE    499 AA; 54852 MW; 1FB7AE50DA01C03F CRC64;
Query Match      100.0%; Score 78; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 1,4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 AYHGWOODIYSL 13
DB      99 AYHGWOODIYSL 111
RESULT 5
ID      AMY2_SCHPO      STANDARD:      PRT:      513 AA.
AC      014154;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      15-JUL-1998 (Rel. 36, Last annotation update)
DE      Probable alpha-amylose C4A8.01 precursor (EC 3.2.1.1) (1,4-alpha-D-
DE      glucan glucanohydrolase).
GN      SPAC4A8.01.

```

```

OS      Schizosaccharomyces pombe (fission yeast).
OC      Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC      Schizosaccharomycetales; Schizosaccharomycetaceae;
OC      Schizosaccharomyces.
OX      NCBI_TaxID=4896;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=972;
RA      Skellton J., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL      Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC      -1 CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC      linkages in oligosaccharides and polysaccharides.
CC      -1 COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
CC      SIMILARITY).
CC      -1 SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC      KNOWN AS THE ALPHA-AMYLASE FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL: Z98762; CAB1471.1; -.
DR      HSSP: P10529; 7TAA.
DR      InterPro: IPR000461; Alpha-amylose.
DR      Pfam: PF00128; alpha-amylose; 1.
KW      Hypothetical protein; Hydrolyase; Glycosidase; Carbohydrate metabolism;
KW      Calcium; Glycoprotein; Signal.
FT      SIGNAL      1
FT      CHAIN      26 513
FT      ACT_SITE    226 226
FT      ACT_SITE    230 230
FT      ACT_SITE    250 250
FT      ACT_SITE    318 318
FT      ACT_SITE    52 60
FT      DISULFID    171 184
FT      DISULFID    260 304
FT      DISULFID    454 488
FT      DISULFID    162 162
FT      CARBOHYD    357 357
SQ      SEQUENCE    513 AA; 58715 MW; 455DD9FEA428C182 CRC64;
Query Match      85.9%; Score 67; DB 1; Length 513;
Best Local Similarity 76.9%; Pred. No. 0.00096;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY      1 AYHGWOODIYSL 13
DB      100 AYHGWOODIYSL 112
RESULT 6
ID      AMYA_ASPEG      STANDARD:      PRT:      484 AA.
AC      P56271;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      15-JUL-1998 (Rel. 36, Last annotation update)
DE      Acid alpha-amylose (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
DE      Aspergillus niger.
OS      Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC      Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX      NCBI_TaxID=5061;
RN      [1]
RP      X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX      MEDLINE=91002514; PubMed=2207069;
RA      Boel E., Brady L., Brzozowski A.M., Derewenda Z., Dodson G.G.,
RA      Jensen V.J., Petersen S.B., Swift H., Thim L., Woldike H.F.;
RT      "Calcium binding in alpha-amyloses: an X-ray diffraction study at

```


RT 2.1-A resolution of two enzymes from *Aspergillus*.
 RL Biochemistry 29:6244-6249(1990).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.
 DR PDB: 2AAA; 15-JUL-93.
 DR InterPro: IPR000461; Alpha-amylase.
 KW Pfam: PF00128; alpha-amylase; 1.
 DR Hydrolyase: Glycosidase; Carbohydrate metabolism; Calcium; KM Glycoprotein; 3D-structure.
 FT ACT_SITE 206 206
 FT ACT_SITE 210 210
 FT ACT_SITE 230 230
 FT ACT_SITE 297 297
 FT DISULFID 30 38
 FT DISULFID 150 164
 FT DISULFID 240 283
 FT DISULFID 440 475
 FT CARBOHYD 24 24
 FT CARBOHYD 157 157
 FT CARBOHYD 197 197
 FT SEQUENCE 484 AA; 52935 MW; 04D596E34680656D CRC64;

Query Match 82.1%; Score 64; DB 1; Length 484;
 Best Local Similarity 76.9%; Pred. No. 0.0028;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVHGYWQODIYSL 13
 ||||| :|| :
 Db 78 AVHGYWQKIDYV 90

RESULT 7
 AML1_SACFI STANDARD: PRT; 494 AA.
 AC P21567;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Alpha-amylase precursor (EC 3.2.1.1).
 GN ALP1.
 OS Saccharomycopsis fibuligera (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
 OX NCBI_TaxID=4944;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87276512; PubMed=3497057;
 RA Itoh T., Yamashita I., Fukui S.;
 RT "Nucleotide sequence of the alpha-amylase gene (ALP1) in the yeast *Saccharomycopsis fibuligera*."
 RL FEBS Lett. 219:339-342(1987).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: X05791; CAA29233.1; .
 DR PIR: S00064; ALBYAF.
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.

DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02806; alpha-amylase.C; 1.
 KW Hydrolyase: Glycosidase; Carbohydrate metabolism; Calcium; KM Glycoprotein; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 494
 FT ACT_SITE 233 233
 FT ACT_SITE 237 237
 FT ACT_SITE 324 324
 FT CARBOHYD 224 224
 FT DISULFID 57 65
 FT DISULFID 177 191
 FT DISULFID 267 310
 FT DISULFID 462 493
 FT SEQUENCE 494 AA; 54387 MW; 7E7DBFDB58B67 CRC64;

Query Match 74.4%; Score 58; DB 1; Length 494;
 Best Local Similarity 61.5%; Pred. No. 0.028;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 AVHGYWQODIYSL 13
 ||||| :|| :
 Db 105 AVHGYWKNRYKI 117

RESULT 8
 YDYL_SCHPO STANDARD: PRT; 478 AA.
 AC Q10427; Q9UNNO;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 55.4 kDa protein C18.01c in chromosome III.
 GN SPEC188.01c OR SPEC11E10.09c.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=972.
 RA Ramsperger U., Pohl T., Wood V., Rajandream M.A., Barrell B.G.;
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: AL121783; CAB57851.1; .
 DR EMBL: AL049662; CAB41221.1; .
 DR HSSP: P10529; 7TAA.
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Hypothetical protein: Hydrolase; Glycosidase.
 FT ACT_SITE 216 216
 FT ACT_SITE 220 220
 FT ACT_SITE 308 308
 FT SEQUENCE 478 AA; 55442 MW; 3C65EAD005DF7D3E CRC64;

Query Match 71.8%; Score 56; DB 1; Length 478;
 Best Local Similarity 69.2%; Pred. No. 0.058;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AVHGYWQODIYSL 13

Db 89 AYHGWMQDMTOL 101

```

RESULT 9
AMYL_SCHPO STANDARD: PRT: 581 AA.
ID AMYL_SCHPO
AC Q09840:
DE 01-FEB-1996 (Rel. 33, Created)
DE 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-FEB-1996 (Rel. 33, Last annotation update)
DE Probable alpha-amylase C23D3.14C precursor (EC 3.2.1.1) (1,4-alpha-D-
DE glucan glucanohydrolase).
DE SPAC23D3.14C.
GN Schizosaccharomyces pombe (Fission yeast).
OS Schizosaccharomycetes; Ascomycota; Schizosaccharomycetes;
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC Schizosaccharomycetes; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=972;
RA Niblett D., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: Z64354; CAA91249.1; -.
CC DR HSSP: P10529; 7TAA.
CC DR InterPro: IPR000461; Alpha-amylase.
CC DR Pfam: PF00128; alpha-amylase; 1.
CC KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
CC Calcium; Glycoprotein; Signal.
CC FT SIGNAL 1 24
CC FT ACT_SITE 233 233
CC FT ACT_SITE 237 237
CC FT ACT_SITE 325 325
CC FT DISULFID 176 64
CC FT DISULFID 267 311
CC FT CARBOHYD 291 291
CC FT CARBOHYD 332 332
CC FT SIGNAL 581 AA; 67004 MW; FE9DE9D323E1890 CRC64;
SQ

```

Query Match 66.78; Score 52; DB 1; Length 581;
 Best local similarity 53.84; Pred. No. 0.33; 3; Indels 0; Gaps 0;
 Matches 7; Conservative 3; Mismatches 3;

QY 1 AYHGWMQDMTOL 13
 DB 104 SYHGWMQDMTOL 116

RESULT 10
 AMYL_DEBOC STANDARD: PRT: 512 AA.
 AC P19269;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amylase 1 precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
GN AMYL.
OS Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=27300;
RN [1]
RN SEQUENCE FROM N.A.
RN STRAIN=ATCC 26077;
RX MEDLINE=90032659; PubMed=2806251;
RA Strasser A.W.M., Selk R., Dolmen R.J., Niermann T., Bielefeld M.,
RA Seeboth P., Tu G., Hollenberg C.P.;
RT "Analysis of the alpha-amylase gene of Schwanniomyces occidentalis
RT and the secretion of its gene product in transformants of different
RT yeast genera."
RT Eur. J. Biochem. 184:699-706(1989).
RL [2]
RN SEQUENCE FROM N.A.
RN STRAIN=CCRC 21164;
RX MEDLINE=92120467; PubMed=1769525;
RA Wu F.M., Wang T.T., Hsu W.H.;
RT "The nucleotide sequence of Schwanniomyces occidentalis alpha-amylase
RL gene."
RL FEBS Microbiol. Lett. 66:313-318(1991).
RN [3]
RN SEQUENCE FROM N.A.
RN STRAIN=ATCC 26077 / CBS 2863;
RX MEDLINE=92307400; PubMed=1612414;
RA Park J.C., Bai S., Tai C.Y., Chun S.B.;
RT "Nucleotide sequence of the extracellular alpha-amylase gene in the
RT yeast Schwanniomyces occidentalis ATCC 26077."
RL FEBS Microbiol. Lett. 72:17-23(1992).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- ENZYME REGULATION: ALPHA-AMYLASE EXPRESSION UNDERLIES CATABOLITE
CC REPRESSION BY GLUCOSE.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: S77586; AAB21151.2; -.
CC DR EMBL: X16040; CAA34162.1; -.
CC DR EMBL: X62079; CAA43995.1; -.
CC DR EMBL: S38381; AAB2383.2; -.
CC DR PIR: S06115; S06115.
CC DR PIR: S23355; S23355.
CC DR HSSP: P10529; 7TAA.
CC DR InterPro: IPR000461; Alpha-amylase.
CC DR Pfam: PF00128; alpha-amylase; 1.
CC KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
CC Glycoprotein; Signal.
CC FT SIGNAL 1 25
CC FT ACT_SITE 242 242
CC FT ACT_SITE 246 246
CC FT ACT_SITE 333 333
CC FT CARBOHYD 233 233
CC FT DISULFID 66 74
CC FT DISULFID 186 200
CC FT DISULFID 276 319
CC FT DISULFID 475 510
CC FT VARIANT 32 32
CC FT VARIANT 26077).

```

```

FT VARIANT 36 36 S -> G (IN STRAIN CCRC 21164).
FT VARIANT 73 73 Y -> I (IN STRAIN ATCC 26077).
FT VARIANT 280 280 N -> S (IN STRAIN CCRC 21164).
FT VARIANT 350 350 D -> A (IN STRAINS CCRC 21164 AND ATCC
FT VARIANT 479 479 26077).
FT VARIANT 479 479 L -> S (IN STRAINS CCRC 21164 AND ATCC
FT VARIANT 483 483 26077).
FT VARIANT 483 483 S -> F (IN STRAIN CCRC 21164).
SO SEQUENCE 512 AA; 56527 MW; 857552B2CF60F965 CRC64;

```

```

Query Match 64.18; Score 50; DB 1; Length 512;
Best Local Similarity 61.58; Pred. No. 0.61;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 1 AYHGYWOODIYSL 13
| | | | | : : : |
Db 114 AYHGYWAKNIDEL 126

```

```

RESULT 11
AMY_BACCI STANDARD; PRT; 528 AA.
ID AMY_BACCI
AC P08137;
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amyglase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
DE Bacillus circulans.
OS Bacillus/Firmicutes; Bacillus/Clostridium group;
OC Bacillus/staphylococcus group; Bacillus.
OX NCBI_TaxID=1397;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87246076; PubMed=3109866;
RX Nishizawa M., Ozawa F., Hishinuma F.;
RT "Molecular cloning of an amyglase gene of Bacillus circulans.";
RL DNA 6,255-265(1987).
CC -1 CATALYTIC ACTIVITY: Endohydrolisis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1 SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC
CC EMBL: M16657; AAA22229.1; -
CC PIR: A29083; ALBSU.
CC HSSP: P26827; ICIU.
CC InterPro: IPR000461; Alpha-amyglase.
CC Pfam: PF00128; alpha-amyglase_C; 1.
CC HydroLase; Glycosidase; Carbohydrate metabolism; Signal.
CC SIGNAL 1 28
CC CHAIN 29 528 POTENTIAL.
CC ACT_SITE 258 258 ALPHA-AMYLASE.
CC ACT_SITE 262 262 BY SIMILARITY.
CC ACT_SITE 357 357 BY SIMILARITY.
CC SEQUENCE 528 AA; 57940 MW; 45574020FA29753 CRC64;

```

```

Query Match 62.88; Score 49; DB 1; Length 528;
Best Local Similarity 77.88; Pred. No. 0.92;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 AYHGYWOOD 9
| | | | | : |

```

```

Db 125 AYHGYWARD 133
RESULT 12
CDGT_BAC3 STANDARD; PRT; 712 AA.
ID CDGT_BAC3
AC P09121;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Cyclomaltoextrin glucanotransferase precursor (EC 2.4.1.19)
DE (Cyclodextrin-glycosyltransferase) (Cgtase).
GN Cgt.
OS Bacillus sp. (strain 38-2).
OC Bacillus/Firmicutes; Bacillus/Clostridium group;
OC Bacillus/staphylococcus group; Bacillus.
OX NCBI_TaxID=1412;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-44.
RX MEDLINE=89036108; PubMed=2972812;
RA Kaneo T., Hamamoto T., Horikoshi K.;
RT "Molecular cloning and nucleotide sequence of the cyclomaltoextrin
RT glucanotransferase gene from the alkalophilic Bacillus sp. strain no.
RT 38-2.";
RL J. Gen. Microbiol. 134:97-105(1988).
RN [2]
RP SEQUENCE OF 1-586 FROM N.A.
RA Hamamoto T., Kaneo T., Horikoshi K.;
RT "Nucleotide sequence of the cyclomaltoextrin glucanotransferase
RT (Cgtase) gene from alkalophilic Bacillus sp. strain No. 38-2.";
RL Agric. Biol. Chem. 51:2019-2022(1987).
CC -1 CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation
CC of a 1,4-alpha-D-glucosidic bond.
CC -1 COPFACTOR: BINDS TWO CALCIUM IONS.
CC -1 SUBUNIT: MONOMER.
CC -1 MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE
CC IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND
CC IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER
CC ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN
CC ALPHA-1,4-GLUCOSIDIC LINKAGE FOR CYCLIZING THE
CC MALTOOLIGOSACCHARIDE PRODUCED.
CC -1 SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC
CC EMBL: M19880; AAA22309.1; -
CC EMBL: D00129; BAA00077.1; -
CC PIR: S24193; ALBSG3.
CC HSSP: P05616; IPRM.
CC InterPro: IPR000461; Alpha-amyglase.
CC InterPro: IPR002044; CBD_4.
CC InterPro: IPR002909; IPT_TIG.
CC Pfam: PF00128; alpha-amyglase; 1.
CC Pfam: PF02806; alpha-amyglase_C; 1.
CC Pfam: PF00686; CBD_4; 1.
CC Pfam: PF01833; TIG_1.
CC ProDom: PD001568; CBD_4; 1.
CC Transferrase; Glycosyltransferase; Calcium; Signal.
CC SIGNAL 1 27
CC CHAIN 28 712 CYCLOMALTODEXTRIN GLUCANOTRANSFERASE.
CC DOMAIN 28 165 A1.
CC DOMAIN 166 229 B.
CC DOMAIN 230 433 A2.
CC DOMAIN 434 522 C.
CC DOMAIN 523 608 D.
CC DOMAIN 609 712 E.

```

FT DISUFLD 70 77 BY SIMILARITY.
 FT ACT_SITE 256 256 BY SIMILARITY.
 FT ACT_SITE 284 284 BY SIMILARITY.
 FT ACT_SITE 355 355 BY SIMILARITY.
 FT CONFLICT 582 586 VPGI -> SWRL (IN REF. 2).
 SO SEQUENCE 712 AA: 78249 MW: 4FAABF70BF81BF9 CRC64;

Query Match 62.8%; Score 49; DB 1; Length 712;
 Best Local Similarity 77.8%; Pred. No. 1.3;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGYWOOD 9
 DB 123 AYHGYWARD 131

RESULT 13

AMR_BACS8 STANDARD; PRT: 713 AA.

AC P17692;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Raw-starch-digesting amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase)
 OS Bacillus sp. (Strain B1018).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1417;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-47.
 RX MEDLINE=90147765; PubMed=1689153;
 RA Riktor P., Tskagatos N., Uda S.;
 RT "Nucleotide sequence of the raw-starch-digesting amylase gene from
 RT Bacillus sp. B1018 and its strong homology to the cyclodextrin
 RT glucanotransferase genes";
 RL Biochem. Biophys. Res. Commun. 166:630-636(1990).
 CC -1- FUNCTION: THIS ENDO-TYPE ADSORBABLE AMYLASE IS CAPABLE TO
 CC DIEST RAW-STARCH.
 CC -1- CATALYTIC ACTIVITY: Endohydrolays of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: M33302; AAA2239.1; -;
 DR EMBL: D90112; BAA14140.1; -;
 DR PIR: A34648; A34648.
 DR PIR: S09196; S09196.
 DR HSSP: P43379; ICDG.
 DR InterPro: IPR000461; Alpha-amylase.
 DR InterPro: IPR002044; CBD_4.
 DR InterPro: IPR002909; IPT_TIG.
 DR Pfam: PF00128; alpha-amylase_1.
 DR Pfam: PF02806; alpha-amylase_C; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR Pfam: PF01833; TIG; 1.
 DR Prodom: PD001568; CBD_4; 1.
 KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 713 RAW-STARCH-DIGESTING AMYLASE.
 FT ACT_SITE 256 256 BY SIMILARITY.
 FT ACT_SITE 260 260 BY SIMILARITY.

FT ACT_SITE 355 355 BY SIMILARITY.
 SO SEQUENCE 713 AA: 77420 MW: 85FB616DA67B888 CRC64;

Query Match 62.8%; Score 49; DB 1; Length 713;
 Best Local Similarity 77.8%; Pred. No. 1.3;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGYWOOD 9
 DB 123 AYHGYWARD 131

RESULT 14

CDG2_PAEMA STANDARD; PRT: 713 AA.

AC P31835;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Cyclomalto-dextrin glucanotransferase precursor (EC 2.4.1.19)
 DE (Cyclodextrin-glycosyltransferase) (CGTase).
 OS Paenibacillus macerans (Bacillus macerans).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Paenibacillus.
 OX NCBI_TaxID=44252;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-37.
 RA Sugimoto T., Kubota M., Sakai S.;
 RT "Polypeptide possessing cyclomalto-dextrin glucanotransferase
 RT activity";
 RL Patent number GB2169902, 23-JUL-1986.
 CC -1- CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation
 CC of a 1,4-alpha-D-glucosidic bond.
 CC -1- COFACTOR: BINDS TWO CALCIUM IONS.
 CC -1- SUBUNIT: MONOMER.
 CC -1- MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS. THE ONE
 CC IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GUCCOSIDIC BOND
 CC IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER
 CC ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN
 CC ALPHA-1,4-GUCCOSIDIC LINKAGE FOR CYCLIZING THE
 CC MATTOLOISACCHARIDE PRODUCED.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.

DR PIR: S26589; ALEBXR.
 DR HSSP: P43379; ICDG.
 DR InterPro: IPR000461; Alpha-amylase.
 DR InterPro: IPR002044; CBD_4.
 DR InterPro: IPR002909; IPT_TIG.
 DR Pfam: PF00128; alpha-amylase_1.
 DR Pfam: PF02806; alpha-amylase_C; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR Pfam: PF01833; TIG; 1.
 DR Prodom: PD001568; CBD_4; 1.
 KW Transferase; Glycosyltransferase; Calcium; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 713 CYCLOMALTODEXTRIN GLUCANOTRANSFERASE.
 FT DOMAIN 28 165 A1.
 FT DOMAIN 166 229 B.
 FT DOMAIN 230 434 A2.
 FT DOMAIN 435 522 C.
 FT DOMAIN 523 609 D.
 FT DOMAIN 610 713 E.
 FT ACT_SITE 256 256 BY SIMILARITY.
 FT ACT_SITE 285 285 BY SIMILARITY.
 FT ACT_SITE 356 356 BY SIMILARITY.
 SO SEQUENCE 713 AA: 76857 MW: 5A287BCCA4AE635 CRC64;

Query Match 62.8%; Score 49; DB 1; Length 713;
 Best Local Similarity 77.8%; Pred. No. 1.3;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGYWOOD 9
 Db 123 AYHGYWPRD 131

RESULT 15

CGGT_BACSO
 ID CDGT_BACSO STANDARD; PRT: 713 AA.
 AC P05618;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Cyclomaltodextrin glucanotransferase precursor (EC 2.4.1.19)
 GN (Cyclodextrin-glycosyltransferase) (Cgtase).
 OS Bacillus sp. (strain 1011).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1410;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87308036; PubMed-2957361;
 RA Kimura K., Kataoka S., Ishii Y., Takano T., Yamane K.;
 RT "Nucleotide sequence of the beta-cyclodextrin glucanotransferase gene
 RT of alkalophilic Bacillus sp. strain 1011 and similarity of its amino
 RT acid sequence to those of alpha-amylases.";
 RL J. Bacteriol. 169:4399-4402(1987).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 RA Harata K., Haga K., Nakamura A., Aoyagi M., Yamane K.;
 RT "X-ray structure of cyclodextrin glucanotransferase from alkalophilic
 RT Bacillus SP. 1011. Comparison of two independent molecules at 1.8-A
 RT resolution.";
 RL Acta Crystallogr. D 52:1136-1145(1996).
 CC -1- CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation
 CC of a 1,4-alpha-D-glucosidic bond.
 CC -1- COFACTOR: BINDS TWO CALCIUM IONS.
 CC -1- SUBUNIT: MONOMER.
 CC -1- MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE
 CC IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND
 CC IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER
 CC ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN
 CC ALPHA-1,4-GLUCOSIDIC LINKAGE FOR CYCLIZING THE
 CC MALTOOLIGOSACCHARIDE PRODUCED.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M17366; AAA2308.1; -.
 DR PIR: A26678; ALBSG1.
 DR PDB: 1PAM; 1JAN-97.
 DR InterPro: IPR000461; Alpha-amylase.
 DR InterPro: IPR002044; CBD_4.
 DR InterPro: IPR002909; IPT_TIG.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF02806; alpha-amylase_C; 1.
 DR Pfam: PF00886; CBD_4; 1.
 DR Pfam: PF01833; TIG; 1.
 DR Pfam: PD001568; CBD_4; 1.
 DR Transferrase; Glycosyltransferase; Calcium; Signal; 3D-structure.
 KW SIGNAL
 FT CHAIN 1 27
 FT SIGNAL 1 27
 FT CHAIN 28 713 CYCLOMALTODEXTRIN GLUCANOTRANSFERASE.
 FT DOMAIN 28 165 A1.
 FT DOMAIN 166 229 B.
 FT DOMAIN 230 433 A2.
 FT DOMAIN 434 522 C.

FT DOMAIN 523 609 D.
 FT DOMAIN 610 713 E.
 FT DISULFD 70 77 BY SIMILARITY.
 FT ACT_SITE 256 256 BY SIMILARITY.
 FT ACT_SITE 284 284 BY SIMILARITY.
 FT ACT_SITE 355 355 BY SIMILARITY.
 SQ SEQUENCE 713 AA; 78340 MW; 524B259526B56C52 CRC64;

Query Match 62.8%; Score 49; DB 1; Length 713;
 Best Local Similarity 77.8%; Pred. No. 1.3;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 AYHGYWOOD 9
 Db 123 AYHGYWARD 131

Search completed: June 13, 2002, 09:11:49
 Job time: 1029 sec

Gene 79, 107-117, 1989
 A>Title: Aspergillus oryzae has two nearly identical Taka-amylose genes, each containing
 A:Reference number: JS0240; MUID:89378767
 A:Accession: JS0240
 A:Molecule type: DNA
 A:Residues: 1-499 <GEN>
 A:Genetics: AMY2
 A>Note: The authors refer to this as isozyme II
 R:Isomura, S.; Ikenaka, T.
 J. Biochem. 74, 1-10, 1973
 A:Reference number: A91930; MUID:74001521
 A:Accession: A91930
 A:Molecule type: protein
 A:Residues: 206-225 <ISE>
 R:Narita, K.
 Proc. Jpn. Acad. 51, 285-290, 1975
 A:Reference number: A93767
 A:Accession: A93767
 A:Molecule type: protein
 A:Residues: 434-443,446-447, 'Q',449-458, 'GTRV',459-464,467-468, 'B',470, 'B',472-499 <NMR>
 R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
 J. Biochem. 95, 697-702, 1984
 A>Title: Structure and possible catalytic residues of Taka-amylose A.
 A:Reference number: A37454; MUID:84212370
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms
 R:Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkmenburg, J.P.; Wilkinson, A.
 submitted to the Brookhaven Protein Data Bank, August 1992
 A:Reference number: A51548; PDB:67HA
 A:Contents: annotation; X-ray crystallography, 2.1 angstroms; residues 22-497
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics: <AMY1>
 A:Gene: amy1
 A:introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Genetics: <AMY2>
 A:Gene: amy2; AMY11
 A:introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; polys
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:122-499/Product: alpha-amylose 1 #status experimental <MAT>
 F:194-321/Domain: alpha-amylose core homology <AMY>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 78; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
 ||||||||||||
 DB 99 AYHGWOODIYSL 111

RESULT 3
 ALAS3
 alpha-amylose (EC 3.2.1.1) 3 precursor - Aspergillus oryzae
 N:Alternate names: alpha-amylose isozyme I; glycogenase; Taka-amylose A
 C:Species: Aspergillus oryzae
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
 C:Accession: S04549; A33215; A44713
 R:Witresl, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989
 A>Title: Three alpha-amylose genes of Aspergillus oryzae exhibit identical intron-exon
 A:Reference number: S04548; MUID:89237897
 A:Accession: S04549
 A:Molecule type: DNA
 A:Residues: 1-499 <MIR>

A:Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 A:Accession: A33215
 A>Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-499 <MT2>
 A:Cross-references: GB:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 R:Genes, M.J.; Dove, M.; Sellig, V.L.
 Gene 79, 107-117, 1989
 A>Title: Aspergillus oryzae has two nearly identical Taka-amylose genes, each contain
 A:Reference number: JS0240; MUID:89378767
 A:Accession: A44713
 A:Molecule type: DNA
 A:Residues: 1-499 <GEN>
 A>Note: The authors refer to this as isozyme I
 R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
 J. Biochem. 95, 697-702, 1984
 A>Title: Structure and possible catalytic residues of Taka-amylose A.
 A:Reference number: A37454; MUID:84212370
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms
 C:Comment: One atom of calcium per molecule is essential for activity.
 C:Genetics:
 A:Gene: amy1
 A:introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
 C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:122-499/Product: alpha-amylose 3 #status experimental <MAT>
 F:194-321/Domain: alpha-amylose core homology <AMY>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 78; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
 ||||||||||||
 DB 99 AYHGWOODIYSL 111

RESULT 4
 B48305
 alpha-amylose (EC 3.2.1.1) B precursor - Aspergillus awamori
 C:Species: Aspergillus awamori
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
 C:Accession: B48305
 R:Kotman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J
 Curr. Genet. 17, 203-212, 1990
 A>Title: Cloning, characterization, and expression of two alpha-amylose genes from As
 A:Reference number: A48305; MUID:90254827
 A:Accession: B48305
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-499 <KOR>
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:194-321/Domain: alpha-amylose core homology <AMY>

Query Match 100.0%; Score 78; DB 2; Length 499;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13

Db 99 AYHGWOODIYSL 111
|||||

RESULT 5

alpha-amylose (EC 3.2.1.1) precursor - Aspergillus sp.
C:Species: Aspergillus sp.
C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
C:Accession: J50663
R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
Biosci. Biotechnol. Biochem. 56, 174-179, 1992.
A:Title: Cloning of the alpha-amylose cDNA of Aspergillus shirosamii and its expression
A:Reference number: J50663; MUID:92323146
A:Accession: J50663
A:Molecule type: mRNA
A:Residues: 1499 <SHI>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylose #status predicted <ALP>
F:194-321/Domain: alpha-amylose core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 78; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
|||||

Db 99 AYHGWOODIYSL 111

RESULT 6

alpha-amylose (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: glycogenase; Taka-amylose A
C:Species: Aspergillus oryzae
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: J70466
R:Tada, S.; Iimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
Agric. Biol. Chem. 53, 593-599, 1989
A:Title: Cloning and nucleotide sequence of the genomic Taka-amylose A gene of Aspergillus
A:Reference number: J70466
A:Accession: J70466
A:Molecule type: DNA
A:Residues: 1499 <TAD>
C:Comment: See also PIR:JK0201 and PIR:J50240.
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics:
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylose #status predicted <AMY>
F:194-321/Domain: alpha-amylose core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 78; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
|||||

Db 99 AYHGWOODIYSL 111

RESULT 7

alpha-amylose (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: Taka-amylose A
C:Species: Aspergillus oryzae
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
C:Accession: JN0588
R:Tsukagoshi, N.; Furukawa, M.; Nagaba, H.; Kiritani, N.; Tsuboi, A.; Udaoka, S.
Gene 84, 319-327, 1989
A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylose A: evidence for
A:Reference number: JN0588; MUID:90128276
A:Accession: JN0588
A:Molecule type: mRNA
A:Residues: 1499 <TSU>
C:Comment: The alpha amyloses are encoded by multigene family.
C:Genetics:
A:Gene: Taa-G1
A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylose #status predicted <AMY>
F:194-321/Domain: alpha-amylose core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 78; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
|||||

Db 99 AYHGWOODIYSL 111

RESULT 8

alpha-amylose precursor - fission yeast (Schizosaccharomyces pombe)
T38770
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
C:Accession: T38770
R:Skellton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21751
A:Accession: T38770
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1513 <SKE>
A:Cross-references: EMBL:Z98762; PIDN:CA81471.1; GSPDB:GN00066; SPDB:SPAC4A8.01
A:Experimental source: strain 972h-; cosmid c4A8
C:Genetics:
A:Gene: SPDB:SPAC4A8.01
A:Map position: 1
C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology

Query Match 85.9%; Score 67; DB 2; Length 513;
Best Local Similarity 76.9%; Pred. No. 0.0024;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AYHGWOODIYSL 13
|||||

Db 100 AYHGWOODIYSL 112

RESULT 9

S72270

alpha-amylase (EC 3.2.1.1) precursor - *Cryptococcus* sp. (strain CS2)
 C:Species: *Cryptococcus* sp.
 A:Variety: strain CS2
 C>Date: 23-Apr-1998 #sequence_revision 08-May-1998 #text_change 20-Jun-2000
 C:Accession: S72270
 R:RefSeq: H.; Chino, M.; Kato, M.; Iimura, Y.
 Biochem. J. 318, 989-996, 1996
 A>Title: Raw-starch-digesting and thermostable alpha-amylase from the yeast *Cryptococcus*
 A:Reference number: S72270; MUID:96433120
 A:Accession: S72270
 A:Molecule type: DNA
 A:Residues: 1-631 <IEF>
 A:Cross-references: EMBL:D83540; NID:g1595852; PIDN:BA12010.1; PID:g1595853
 A:Experimental source: strain S-2
 C:Genetics:
 A:Gene: amy-CS2
 A:Insertions: 289/2; 326/2
 C:Function:
 A:Pathway: glycogen/starch degradation
 C:Superfamily: *Aspergillus* alpha-amylase; alpha-amylase core homology
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-631/Product: alpha-amylase #status predicted <MAT>
 F:206-335/Domain: alpha-amylase core homology <AMY>

Query Match 84.6%; Score 66; DB 2; Length 631;
 Best Local Similarity 76.9%; Pred. No. 0.0044;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
 ||||| |||||

DB 105 AYHGWOODIYEL 117

RESULT 10

alpha-amylase (EC 3.2.1.1) - *Aspergillus niger*
 C:Species: *Aspergillus niger*
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A35282
 R:Boel, E.; Brady, L.; Brzozowski, A.M.; Derewenda, Z.; Dodson, G.G.; Jensen, V.J.; Pete
 Biochemistry 29, 6244-6249, 1990
 A>Title: Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-angstrom
 A:Reference number: A35282; MUID:91002514
 A:Accession: A35282
 A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: mRNA
 A:Residues: 1-484 <BOE>
 C:Function:
 A>Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: *Aspergillus* alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolyase; polysaccharide degradation
 F:173-300/Domain: alpha-amylase core homology <AMY>

Query Match 82.1%; Score 64; DB 1; Length 484;
 Best Local Similarity 76.9%; Pred. No. 0.0071;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
 ||||| |||||

DB 78 AYHGWOODIYDV 90

RESULT 11

alpha-amylase (EC 3.2.1.1) SWA2 precursor - yeast (*Schwannomyces occidentalis*)
 N:Alternate names: alpha-1,4 glucanohydrolyase
 C:Species: *Schwannomyces occidentalis*
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: S33921

R:Claros, M.G.; Abarca, D.; Fernandez-Lobato, M.; Jimenez, A.
 Curr. Genet. 24, 75-83, 1993
 A>Title: Molecular structure of the SWA2 gene encoding an AMY1-related alpha-amylase
 A:Reference number: S33921; MUID:93365041
 A:Accession: S33921
 A:Molecule type: DNA
 A:Residues: 1-507 <CLIA>
 A:Cross-references: EMBL:X73497; NID:g396561; PIDN:CA51912.1; PID:g396562
 C:Genetics:
 A:Gene: SWA2
 C:Function:
 A>Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: *Aspergillus* alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolyase; polysaccharide degradation
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-507/Product: alpha-amylase #status predicted <MAT>
 F:205-332/Domain: alpha-amylase core homology <AMY>
 F:134,229/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.1%; Score 64; DB 2; Length 507;
 Best Local Similarity 69.2%; Pred. No. 0.0075;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
 ||||| |||||

DB 110 AYHGWOODIYAI 122

RESULT 12

alpha-amylase (EC 3.2.1.1) - *Aspergillus oryzae*
 N:Alternate names: glycogenase; Taka-amylase A
 C:Species: *Aspergillus oryzae*
 C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
 C:Accession: JK0201
 R:Roda, H.; Kondo, K.; Narita, K.
 Proc. Jpn. Acad. 58B, 208-212, 1982
 A>Title: The complete amino acid sequence of Taka-amylase A.
 A:Reference number: JK0201
 A:Accession: JK0201
 A:Molecule type: protein
 A:Residues: 1-478 <IOD>
 C:Comment: One atom of calcium per molecule is essential for the activity.
 C:Comment: This enzyme is a glycoprotein.
 C:Comment: See also PIR:JT0466 and PIR:J50240.
 C:Function:
 A>Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: *Aspergillus* alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolyase; polysaccharide degradation
 F:173-300/Domain: alpha-amylase core homology <AMY>
 F:197/binding site: carbohydrate (Asn) (covalent) #status experimental
 F:210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 79.5%; Score 62; DB 2; Length 478;
 Best Local Similarity 84.6%; Pred. No. 0.015;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AYHGWOODIYSL 13
 ||||| |||||

DB 78 AYTGWQTDIYSL 90

RESULT 13

alpha-amylase (EC 3.2.1.1) precursor - yeast (*Lipomyces kononenkoae*)
 N:Alternate names: LKA1 protein; raw starch-degrading amylase
 N:Contains: alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41)
 C:Species: *Lipomyces kononenkoae*
 C>Date: 07-Feb-1996 #sequence_revision 23-Aug-1996 #text_change 18-Jun-1999

C:Accession: J04510; PC4116
 R:Steyn, A.J.C.; Marmur, J.; Pretorius, I.S.
 Gene 166, 65-71, 1995
 A:Title: Cloning, sequence analysis and expression in yeasts of a cDNA containing a lipase
 A:Reference number: J04510; M01D:96105202
 A:Accession: J04510
 A:Molecule type: mRNA
 A:Residues: 1-624 <ST2>
 A:Cross-references: GB:030376; NID:g1173536; PIDN:AAC49622.1; PID:g1173537
 A:Experimental source: strain IGC4052B
 A:Accession: PC4116
 A:Molecule type: protein
 A:Residues: 29-44 <ST2>
 A:Experimental source: IGC4052B
 C:Genetics:
 A:Gene: LKAL
 C:Function:
 A:Description: hydrolyzes raw starch, degrading both alpha-1,4 and alpha-1,6 linkages
 A:Pathway: glycogen/starch degradation
 C:Superfamily: lipomyc alpha-amylase; alpha-amylase core homology; glucoamylase starch
 C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide
 F:1-28/Domain: signal sequence #status predicted <SIG>
 F:29-624/Product: alpha-amylase #status predicted <MAT>
 F:48-141/Domain: glucoamylase starch-binding domain homology <SBD>
 F:120-447/Domain: alpha-amylase core homology <AMY>
 F:177-185,297-311,367-430,587-622/Disulfide bonds: #status predicted
 F:504,344/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:357,377,444/Active site: His, Glu, Asp #status predicted

Query Match 79.5%; Score 62; DB 1; Length 624;
 Best Local Similarity 69.2%; Pred. No. 0.02;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AYHGWOODIYSL 13
 DB 225 AYHGWMKDIKFL 237

RESULT 14
 S31478
 alpha-amylase (EC 3.2.1.1) - Thermactinomyces vulgaris
 C:Species: Thermactinomyces vulgaris
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 22-Oct-1999
 C:Accession: S31478
 R:Hoefmeister, B.; Koenig, S.; Hoang, V.; Engel, J.; Mayer, G.; Heese, O.; Hansen, G.; H
 submitted to the EMBL Data Library, December 1992
 A:Description: The gene amyTV coding for a non-glucoamylase alpha-amylase from Thermactin
 A:Reference number: S31478
 A:Accession: S31478
 A:Molecule type: DNA
 A:Residues: 1-482 <HO>
 A:Cross-references: EMBL:X69807; NID:g48289; PIDN:CAA49465.1; PID:g48290
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation
 F:194-319/Domain: alpha-amylase core homology <AMY>

Query Match 75.6%; Score 59; DB 2; Length 482;
 Best Local Similarity 69.2%; Pred. No. 0.048;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 AYHGWOODIYSL 13
 DB 107 AYHGWTYDFYSV 119

RESULT 15
 ALBYAF
 alpha-amylase (EC 3.2.1.1) precursor - yeast (Saccharomycopsis fibuligera)

C:Species: Saccharomycopsis fibuligera
 C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 18-Jun-1999
 C:Accession: S00064
 R:Itou, T.; Yamashita, I.; Fukui, S.
 FEBS Lett. 219, 339-342, 1987
 A:Title: Nucleotide sequence of the alpha-amylase gene (ALP1) in the yeast Saccharomy
 A:Reference number: S00064; M01D:87276512
 A:Accession: S00064
 A:Molecule type: DNA
 A:Residues: 1-494 <ITO>
 A:Cross-references: EMBL:X05791; NID:g4847; PIDN:CAA29233.1; PID:g4848
 A>Note: the sequence from Fig. 3 is inconsistent with that from Fig. 1 in having 366-
 C:Genetics:
 A:Gene: ALP1
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: calcium; extracellular protein; glycoprotein; glycosidase; hydrolase; met
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-494/Product: alpha-amylase #status predicted <MAT>
 F:200-327/Domain: alpha-amylase core homology <AMY>
 F:57-65,177-191,267-310,462-493/Disulfide bonds: #status predicted
 F:148,189,202,237/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:224/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:233,257,324/Active site: Asp, Glu, Asp #status predicted

Query Match 74.4%; Score 58; DB 1; Length 494;
 Best Local Similarity 61.5%; Pred. No. 0.072;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AYHGWOODIYSL 13
 DB 105 AYHGWMKNIYKI 117

Search completed: June 13, 2002, 08:55:49
 Job time: 369 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 13, 2002, 08:53:25 ; Search time 158.68 seconds
(without alignments)
9.100 Million cell updates/sec

Title: US-09-710-339-2_COPY_98_110
Perfect score: 78
Sequence: 1 AHGTYQDDIYSL 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	478	15 AAR46065	Mutant alpha-amy
2	78	100.0	478	16 AAR72450	Aspergillus oryzae
3	78	100.0	478	16 AAR78270	Aspergillus oryzae
4	78	100.0	478	17 AAW14500	Aspergillus oryzae
5	78	100.0	498	22 AAB84206	Amino acid sequenc
6	69	88.5	55	17 AAR88213	Alpha-amyase frag
7	69	88.5	493	17 AAR88212	Alpha-amyase. Th
8	62	79.5	478	16 AAR79025	Mature taka-amyas
9	58	74.4	468	13 AAR24136	Alpha-amyase vari
10	58	74.4	494	8 AAR70571	Alpha-amyase gene
11	54	69.2	468	15 AAR63184	Variant alpha amy1

12	53	67.9	468	15 AAR63185	Variant alpha amy1
13	50	64.1	452	15 AAR63186	Variant alpha amy1
14	50	64.1	511	9 AAR61161	Recombinant alpha-
15	50	64.1	511	9 AAR61180	Sequence of alpha-
16	50	64.1	512	11 AAR07574	Alpha-amyase enco
17	49	62.8	468	15 AAR63187	Variant alpha amy1
18	49	62.8	656	11 AAR06109	Sequence of cyclom
19	49	62.8	666	17 AAR61518	Sequence of cyclom
20	49	62.8	666	17 AAR06773	Wild type CGTase.
21	49	62.8	666	17 AAR17592	Bacillus CGTase va
22	49	62.8	666	17 AAR17593	Bacillus CGTase va
23	49	62.8	666	17 AAR17594	Bacillus CGTase va
24	49	62.8	666	17 AAR17595	Bacillus CGTase va
25	49	62.8	666	17 AAR17596	Bacillus CGTase va
26	49	62.8	666	17 AAR17597	Bacillus CGTase va
27	49	62.8	666	17 AAR17598	Bacillus CGTase va
28	49	62.8	666	17 AAR17599	Bacillus CGTase va
29	49	62.8	666	17 AAR17600	Bacillus CGTase va
30	49	62.8	666	17 AAR17602	Bacillus CGTase va
31	49	62.8	666	17 AAR17603	Bacillus CGTase va
32	49	62.8	666	17 AAR17604	Bacillus CGTase va
33	49	62.8	666	17 AAR17605	Bacillus CGTase va
34	49	62.8	666	20 AAR02599	A cyclomalitodextri
35	49	62.8	687	17 AAR17601	Bacillus CGTase va
36	49	62.8	688	17 AAR17568	Bacillus CGTase va
37	49	62.8	688	17 AAR17572	Bacillus CGTase va
38	49	62.8	688	17 AAR17576	Bacillus CGTase va
39	49	62.8	712	12 AAR10051	Cyclomalitodextrin
40	49	62.8	713	11 AAR06110	Sequence of cyclom
41	49	62.8	713	12 AAR10052	Cyclomalitodextrin
42	49	62.8	713	12 AAR12743	Enzyme with starch
43	47	60.3	655	13 AAR20139	Sequence encoded b
44	47	60.3	715	13 AAR20138	Sequence encoded b
45	46	59.0	170	22 AAR25600	Human G Protein-Co

ALIGNMENTS

RESULT 1
AAR46065 standard; protein; 478 AA.
ID AAR46065;
XX AAR46065;
AC
XX 18-JUL-1994 (first entry)
DT
XX
XX Mutant alpha-amyase.
DE
XX Methionine substitution; stability; activity; detergent;
KW dishwashing agents; liquifaction agents.
XX
OS Aspergillus oryzae.
XX
XX
XX W09402597-A.
PN
XX
PD 03-FEB-1994.
XX
XX 06-JUL-1993; 93WO-DK00230.
PF
XX
XX 23-JUL-1992; 92DK-0000946.
PR 16-DEC-1992; 92DK-0001503.
PR 15-MAR-1993; 93DK-0000292.
XX
XX (NOVO) NOVO-NORDISK AS.
PA
XX Bisgard-frantzen H, Svendsen A;
XX
XX WPI, 1994-048855/06.
DR
XX Mutant alpha-amyase from Bacillus species comprising a
PT methionine substitution - with improved stability and activity at
PT low pH, for use in detergents, dishwashing agents and

PT Liquefaction agents
 XX
 PS Claim 1; Page 7; 20pp; English.
 CC The sequence os that of the Asoergillus oryzae alpha amylase, sold
 CC commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can
 CC be mutated by substitution of one or more of its methionine residues
 CC for any amino acid other than cysteine. The mutant alpha-amylase
 CC exhibits a better activity level and better stability in the
 CC presence of oxidizing agents than previous mutant alpha amylases,
 CC and improved thermostability at moderately low pH. The enzyme can
 CC be used as an additive for detergents, dishwashing agents and
 CC liquifaction agents.
 XX
 SQ Sequence 478 AA;

Query Match 100.0%; Score 78; DB 15; Length 478;
 Best Local Similarity 100.0%; Pred. No. 0.00036;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AYHGWMQODIYSL 13
 |||||
 DB 78 ayhgywqgdylsl 90

RESULT 2

AAR72450
 ID AAR72450 standard; Protein: 478 AA.

AC AAR72450;

DT 01-DEC-1995 (first entry)

DE Aspergillus oryzae alpha amylase (mature protein).

KW Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
 KW washing; Bacillus licheniformis; Bacillus amyloliquefaciens;
 KW Bacillus stearothermophilus; dyeing; bleaching; scouring; textile;
 KW thermostable.

OS Aspergillus oryzae.

PN WO9510603-A.

PD 20-APR-1995.

PF 05-OCT-1994; 94WO-DK00370.

PR 08-OCT-1993; 93DK-0001133.

PR 02-FEB-1994; 94DK-0000140.

PA (NOVO) NOVO-NORDISK AS.

PI Svendsen A, Thøgersen M;

PI Van der zee P, Bisgaardtrantzen H, Borchert T;

DR WPI; 1995-161790/21.

PT New Bacillus derived alpha-amylase variants - having amino acid
 PT modifications to improve washing and/or dishwashing performance
 XX
 PS Disclosure; Page 75-76; 105pp; English.

CC Variant alpha amylase enzymes which have improved washing and/or
 CC as detergent additives. The enzymes have one or more amino acid
 CC residues added, deleted or substituted. The variants can also be
 CC used for textile desizing prior to scouring, bleaching and dyeing.
 CC The variants have improved thermostability, acid/alkaline stability;
 CC low temperature optimum; pH optimum; higher hydrolysis velocity and
 CC improved tolerance to other composition constituents, e.g. oxidation
 CC agents.
 CC

SQ Sequence 478 AA;

Query Match 100.0%; Score 78; DB 16; Length 478;
 Best Local Similarity 100.0%; Pred. No. 0.00036;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWMQODIYSL 13
 |||||
 DB 78 ayhgywqgdylsl 90

RESULT 3

AAR78270
 ID AAR78270 standard; Protein: 478 AA.

AC AAR78270;

DT 17-JAN-1996 (first entry)

DE Aspergillus oryzae alpha amylase (mature protein).

KW Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;
 KW starch; thermostable; methionine; Bacillus licheniformis;
 KW Bacillus amyloliquefaciens; Bacillus stearothermophilus;
 KW Aspergillus oryzae.

OS Aspergillus oryzae.

PN WO9521247-A1.

PD 10-AUG-1995.

PF 05-OCT-1994; 94WO-DK00371.

PR 02-FEB-1994; 94DK-0000141.

PA (NOVO) NOVO-NORDISK AS.

PI Marcher D, Nilsson TE, Pedersen HH, Toft AH;

DR WPI; 1995-28367/37.

PT Use of an oxidation stable alpha-amylase - for simultaneous desizing
 PT and bleaching or scouring of fabrics contg. starch or starch derivs.
 XX
 PS Disclosure; Page 25-26; 37pp; English.

CC Oxidation stable alpha amylases can be used for the simultaneous
 CC desizing and bleaching or scouring of a fabric comprising starch or
 CC starch derivatives. They exhibit a better heat stability.
 CC especially in the presence of oxidizing agents. They are obtained
 CC from a parent alpha amylase by replacing one or more methionine
 CC residues with any amino acid different from Cys or Met, preferably
 CC Leu, Thr, Ala, Ser, Ile or Asp. The parent alpha amylase is
 CC pref. derived from a Bacillus species, although alpha amylases of
 CC fungal origin can also be used. This sequence is the wild type
 CC (unmodified) alpha amylase of Aspergillus oryzae.
 CC

SQ Sequence 478 AA;

Query Match 100.0%; Score 78; DB 16; Length 478;
 Best Local Similarity 100.0%; Pred. No. 0.00036;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AYHGWMQODIYSL 13
 |||||
 DB 78 ayhgywqgdylsl 90

RESULT 4

AAM14500

XX	AAW14500:standard; protein; 478 AA.	
AC	AAW14500;	
DT	04-JUN-1997 (first entry)	
XX		
DE	Aspergillus oryzae alpha-amylase (mature protein).	
XX		
KM	alpha-amylase; mature; Bacillus licheniformis; termamyl; fungamyl;	
KM	Aspergillus oryzae; Bacillus amyloliquefaciens; altered property;	
KM	calcium dependency; substrate binding; stability; pH optimum;	
KM	thermostability; cleavage; oligosaccharide substrate; dishwashing;	
KM	washing; detergent additive; fabric desizing; starch liquefaction;	
KM	sweetener; ethanol production; variant.	
XX		
OS	Aspergillus oryzae.	
XX		
PH	Key	Location/Qualifiers
FT	Misc-difference 13..45	/label=
FT		/note=
FT		alpha-amylase (used as a template for a variant)
FT		corresponding to 7-23 of AAW14499 is deleted or
FT		replaced with a fragment corresponding to
FT		this fragment; claim 33"
FT	Misc-difference 14..40	/label=
FT		/note=
FT		alpha-amylase (used as a template for a variant)
FT		corresponding to 12-19 of AAW14499 is deleted or
FT		replaced with a fragment corresponding to
FT		this fragment; claim 30"
FT	Misc-difference 32..38	/label=
FT		/note=
FT		alpha-amylase (used as a template for a variant)
FT		corresponding to 44-57 of AAW14499 is deleted or
FT		replaced with a fragment corresponding to
FT		this fragment; claim 18"
FT	Misc-difference 66..84	/label=
FT		/note=
FT		alpha-amylase (used as a template for a variant)
FT		corresponding to 117-185 of AAW14499 is deleted or
FT		replaced with a fragment corresponding to
FT		this fragment; claim 24"
FT	Misc-difference 70..78	/label=
FT		/note=
FT		alpha-amylase (used as a template for a variant)
FT		corresponding to 117-185 of AAW14499 is deleted or
FT		replaced with a fragment corresponding to
FT		this fragment; claim 24"
FT	Misc-difference 98..120	/label=
FT		/note=
FT		alpha-amylase (used as a template for a variant)
FT		corresponding to 117-185 of AAW14499 is deleted or
FT		replaced with a fragment corresponding to
FT		this fragment; claim 24"
FT	Misc-difference 102..120	/label=
FT		/note=
FT		alpha-amylase (used as a template for a variant)
FT		corresponding to 117-185 of AAW14499 is deleted or
FT		replaced with a fragment corresponding to
FT		this fragment; claim 24"

FT			fragment corresponding to this fragment; claim
FT	26"	Misc-difference 121..181	"at least one amino acid residue of a parent
FT		/note=	alpha-amylase (used as a template for a variant)
FT			corresponding to this fragment is deleted or
FT			replaced with a fragment corresponding to
FT			102-206 of AAM14499; claim 41"
FT	Misc-difference 121..174	/note=	"preferred region where at least one amino acid
FT			residue of a parent alpha-amylase (used as a
FT			template for a variant) corresponding to this
FT			fragment is deleted or replaced with a fragment
FT			corresponding to 102-199 of AAM14499; claim 42"
FT	Misc-difference 165..177	/label=	loop 3 modification region
FT		/note=	"at least one amino acid residue of a parent
FT			alpha-amylase (used as a template for a variant)
FT			corresponding to 195-202 of AAM14499 is deleted
FT			or replaced with a fragment corresponding to
FT			this fragment; claim 21"
FT	Misc-difference 166..173	/label=	loop 3 modification region
FT		/note=	"preferred region where at least one amino acid
FT			residue of a parent alpha-amylase (used as a
FT			template for a variant) corresponding to 196-199
FT			of AAM14499 is deleted or replaced with a
FT			fragment corresponding to this fragment; claim
FT	23"	Misc-difference 181..184	"an amino acid fragment corresponding to this
FT		/note=	region is deleted from the parent sequence of a
FT			variant Fungamyl; claim 43"
FT	Misc-difference 291..313	/label=	loop 8 modification region
FT		/note=	"at least one amino acid residue of a parent
FT			alpha-amylase (used as a template for a variant)
FT			corresponding to 322-346 of AAM14498 is deleted
FT			or replaced with a fragment corresponding to
FT			this fragment; claim 36"
FT	Misc-difference 297..313	/label=	loop 8 modification region
FT		/note=	"preferred region where at least one amino acid
FT			residue of a parent alpha-amylase (used as a
FT			template for a variant) corresponding to 325-346
FT			of AAM14498 is deleted or replaced with a
FT			fragment corresponding to this fragment; claim
FT			38"
XX			
PN	WO9623874-A1.		
XX			
PD	08-AUG-1996.		
XX			
PE	05-FEB-1996;	96WO-DK00057.	
XX			
PR	10-NOV-1995;	95DK-0001256.	
XX	03-FEB-1995;	95DK-0000129.	
PR	23-OCT-1995;	95DK-0001192.	
XX			
PA	(NOVO) NOVO-NORDISK AS.		
XX			
PI	Bisgard-Frantzen H, Borchert TV, Svendsen A;		
XX			
DR	WPI: 1996-371424/37.		
XX			
PT	Alpha-amylase variants and methods of production - have altered		
PT	properties such as calcium dependency, substrate binding and		
PT	stability		
XX			
PS	Disclosure; Page 87-88; 171pp; English.		
XX			
CC	The present sequence is the mature Aspergillus oryzae alpha-amylase (A)		
CC	variants of parent termamyl- and fungamyl-like alpha-amylases (and		
CC	methods of constructing them) are claimed. Examples of variants are		
CC	featured above. The variants have altered properties such as calcium		

CC dependency, substrate binding and stability. Also one or more proline or
 CC cysteine residues in the variant is modified or replaced with a
 CC non-proline or non-cysteine residue such as alanine. The variants can be
 CC used for (dist)washing, as detergent additives or for fabric desizing or
 CC starch liquefaction. They can also be used for the production of
 CC sweeteners and ethanol from starch. See also AAW14498-99.

XX Sequence 478 AA;

Query Match 100.0%; Score 78; DB 17; Length 478;

Best Local Similarity 100.0%; Pred. No. 0.00036;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AYHGWMQDIYSL 13

DB 78 ayhgywqgdlysl 90

RESULT 5

AAB84206
 ID AAB84206 standard; Protein: 498 AA.

XX AAB84206;

DT 06-AUG-2001 (first entry)

DE Amino acid sequence of a fungamyl-like alpha-amylase.

KW Fungamyl-like alpha-amylase; glucosylase; dextrinisation; maltose;

KW alcohol; starch; dough improver; brewing; starch liquification.

OS Aspergillus oryzae.

PN WO200134784-A1.

PD 17-MAY-2001.

PF 10-NOV-2000; 2000WO-DK00625.

PR 10-NOV-1999; 99DK-0001617.

PA (NOVO) NOVOZYMES AS.

PI Bisgard-Frantzen H, Svendsen A, Pedersen S;

DR WPI: 2001-367478/38.

DR N-PSDB: AAF90208.

PT New variant of Fungamyl-like alpha-amylase, useful for production of
 PT maltose syrups, includes mutations that improve stability against heat
 PT and acidic pH -

PS Claim 1; Page 42-45; 49pp; English.

CC The present sequence represents a fungamyl-like alpha-amylase. The
 CC specification describes variants of this fungamyl-like alpha-amylase,
 CC which have an alteration in one the amino acid regions 98-110, 150-160,
 CC 161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion
 CC or substitution of an amino acid or an insertion of an amino acid
 CC downstream of a particular position. The variants retain alpha-amylase
 CC activity, and have better heat stability and/or stability at acidic pH,
 CC relative to wild-type enzyme. The variants can therefore be used at
 CC higher temperatures (more efficient conversion or faster reaction, and
 CC have reduced need for cooling and reduced risk of contamination). The
 CC variants may also be used in conjunction with other enzymes,
 CC particularly glucosylase during dextrinisation. The variants are
 CC used to produce syrups, particularly of high maltose content, or alcohol,
 CC from starch, as dough improver for baked goods, in brewing, to increase
 CC fermentability of the wort, and for liquefaction of starch.

SO Sequence 498 AA;

Query Match 100.0%; Score 78; DB 22; Length 498;

Best Local Similarity 100.0%; Pred. No. 0.00038;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AYHGWMQDIYSL 13

DB 98 ayhgywqgdlysl 110

RESULT 6

AAR88213
 ID AAR88213 standard; Peptide: 55 AA.

XX AAR88213;

DT 03-APR-1996 (first entry)

DE Alpha-amylase fragment.

KW Alpha-amylase; thermostable enzyme; baking;

KW Thermomyces lanuginosus; PCR; polymerase chain reaction.

OS Thermomyces lanuginosus CBS 223.63.

PN WO9601323-A1.

PD 18-JAN-1996.

PF 03-JUL-1995; 95WO-EP02607.

PR 04-JUL-1994; 94GB-0013419.

PA (DANI-) DANISCO AS.

PI Michelsen B, Rasmussen P;

DR WPI: 1996-087673/09.

DR N-PSDB: AAT10565.

PT Thermophilic alpha-amylase with activity range of 60-80 degrees C -
 PT derived from Thermomyces lanuginosus, useful in the prepn. of
 PT foodstuffs and bakery prods. esp. bread

PS Disclosure: Page 40; 94pp; English.

CC Fragments (AAR88213-14) of Thermomyces lanuginosus thermostable

CC alpha-amylase (AAR88212) are the products of DNA fragments (AAT10565-66)

CC generated from T. lanuginosus genomic DNA by PCR using primers based
 CC on conserved regions of alpha-amylase genes.

SO Sequence 55 AA;

Query Match 88.5%; Score 69; DB 17; Length 55;

Best Local Similarity 76.9%; Pred. No. 0.00095;

Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 AYHGWMQDIYSL 13

DB 26 syhgywqgdlysl 38

RESULT 7

AAR88212
 ID AAR88212 standard; Protein: 493 AA.

XX AAR88212;

DT 03-APR-1996 (first entry)

DE Alpha-amylase.

XX 31-OCT-1994; 94JP-0288658.
PF
XX
XX 12-NOV-1993; 93JP-0306096.
XX
PA (EZAKI) EZAKI GLICO CO.
PA (NIDE) NEC CORP.
XX
XX
DR WPI; 1995-279919/37.
XX
PT Modifying a transferase by enhancing hydrophobicity of a selected
PT site - increases transfer activity, also new mutant
PT neo-pullulanase(s)
XX
XX
PS Disclosure; Page 10-11; 18pp; English.
XX
XX This sequence represents the mature form of taka-amylase from *A. oryzae*.
CC This sequence was used in a method for the generation of mutant
CC pullulanases for use in the food industry (see also AAR79026-28). The
CC wild type pullulanase enzyme was modified by the method of the
CC invention for enhancing the hydrophobicity of a selected site of the
CC pullulanase. The method comprises replacement of a group in the
CC selected site with a hydrophobic group, replacement of an amino acid
CC with a hydrophobic amino acid, and/or insertion or deletion of a
CC hydrophobic amino acid from the selected site. The method was used
CC to produce neopullulanases Y377F, S422V and M375L.
XX
XX
XX Sequence 478 AA;

Query Match	79.58;	Score 62;	DB 16;	Length 478;
Best Local Similarity	84.68;	Pred. No. 0.11;		
Matches 11; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	1	AYHGYWQDDIYSL	13
Db	78	aytgywqtdiysl	90

DT	17-NOV-1992 (first entry)
XX	
DE	Alpha-amylase variant with leucine at position 84.

KW Mutant; maltose; malto-oligosaccharides; *Saccharomycopsis*; *fibuligera*; polymerisation; DP; transglycosifier.

	Key	Location/Qualifiers
FH	Misc-difference	84
FT		/note= "mutation to leucine"
FT		

AA JP04108386-A.
PN
XX
PD 09-APR-1992.

XX
PF 28-AUG-1990; 90JP-0226112.

XX
PR 28-AUG-1990; 90JP-0226112.

XX
PA (AGEN) AGENCY OF IND SCI & .

XX
DR WPT: 1992-171652/21.

DR N-PSDB; AAQ25079.
XX

PT Variant alpha-amylase gene for mfr. of malto-oligosaccharide(s) -
is obt'd by mutating the nucleotide at position 329 of the

PT Saccharomycopsis fibuligera wild-type sequence to thymine.
 XX
 PS Claim 2; Fig 1; 10pp; Japanese.
 XX

CC The variant alpha amylase was obtd. from a gene contg. a mutation
 CC of the 329th nucleotide of the alpha amylase gene of Saccharomycopsis
 CC fibuligera to T. This mutation results in substitution of the wild-
 CC type amino acid at position 84 of alpha-amylase by leucine. The
 CC variant alpha-amylase is high in transglycosylating activity.
 CC variant alpha-amylase may be used to prepare malto-oligosaccharides
 CC with a degree of polymerisation (DP) of at least 7, by inversion of
 CC the malto-oligosaccharide.
 XX

SQ Sequence 468 AA;

Query Match 74.4%; Score 58; DB 13; Length 468;
 Best Local Similarity 61.5%; Pred. No. 0.45;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
 |||||:|:|:
 Db 79 ayhgywmknlyki 91

RESULT 10

AA63184 ID AAP70571 standard; Protein; 494 AA.

AC AAP70571;

DT 03-MAY-1991 (first entry)

DE Alpha-amylase gene product from plasmid psl alpha 1.

XX Amylase.

OS Saccharomycetes fibuligera HUT7212.

PN JP62104576-A.

PD 15-MAY-1987.

PF 31-OCT-1985; 85JP-0244892.

PR 31-OCT-1985; 85JP-0244892.

PA (FUKU/) FUKUI S.

DR WPI: 1987-173694/25.

DR N-PSDB; AAN70916.

PT Amylase prodn. - comprises culturing microorganism transformed
 PT with vector deoxyribonucleic acid, accumulating and collecting
 PT amylase
 PS

PS Disclosure; Fig 1; 14pp; Japanese.

CC The product is an amylase biosynthetic enzyme encoded by a plasmid
 CC which may be used to transform an E.coli expression system for
 CC the stable production of amylase, useful in ethanol fermentation.
 XX

SQ Sequence 494 AA;

Query Match 74.4%; Score 58; DB 8; Length 494;
 Best Local Similarity 61.5%; Pred. No. 0.48;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13
 |||||:|:|:
 Db 105 ayhgywmknlyki 117

RESULT 11

AA63184 ID AAR63184 standard; Protein; 468 AA.

AC AAR63184;

DT 16-JUN-1995 (first entry)

DE Variant alpha amylase deriv. from Saccharomycopsis fibuligera.

KM alpha amylase; carbohydrate hydrolase; increased activity;

KM tyrosine residue; enzyme centre; mass production; oligosaccharide;

KM variant; cyclomalto-dextrin glucanotransferase.

XX Saccharomycopsis fibuligera.

XX Key Location/Qualifiers

FT Misc-difference 83 /label= Y83F

FT Misc-difference 238, 256 /note= "this peptide region is missing in the

FT specification, but has been included due to its
 FT similarity to related protein sequences, which
 FT are encoded by nucleotides of exactly the same
 FT sequence."

PN JP06253836-A.

PD 13-SEP-1994.

PF 04-MAR-1993; 93JP-0069303.

PR 04-MAR-1993; 93JP-0069303.

PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.

DR WPI: 1994-328987/41.

DR N-PSDB; AA077665.

PT Variant carbohydrate hydrolase(s) with increased activity -

PT consists of e.g. alpha-amylase with tyrosine residue in enzyme

PS Example 1; Page 13-15; 27pp; Japanese.

CC AAR63184-7 show the amino acid sequences of variant alpha amylases,

CC composed by substituting bases 247-249 of the structural gene (see

CC AA077665-8) region, with TTC, TGG, CTC or AAC. These substitutions

CC result in the 83rd amino acid residue (tyrosine) of the wild type

CC sequence being changed to phenylalanine, tryptophan, leucine or

CC asparagine respectively. The substituted amino acid is present in

CC the active site of the enzyme and confers increased activity on the

CC enzyme. The variants are useful for the mass production of a

CC oligosaccharides. (see AAR63188 for the variant structure of a

CC cyclomalto-dextrin glucanotransferase).

SQ Sequence 468 AA;

Query Match 69.2%; Score 54; DB 15; Length 468;
 Best Local Similarity 53.8%; Pred. No. 1.9;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 AYHGYWQDDIYSL 13

AA63185 ID AAR63185 standard; Protein; 468 AA.

```

AC  AAR63185;
XX
XX  16-JUN-1995 (first entry)
DT
XX  Variant alpha amylase deriv. from Saccharomycopsis fibuligera.
DE
XX  alpha amylase; carbohydrate hydrolase; increased activity;
XX  tyrosine residue; enzyme centre; mass production; oligosaccharide;
XX  variant; cyclomaltohextrin glucanotransferase.
OS  Saccharomycopsis fibuligera.
FH  Key
FT  MISC-difference 83
FT  /label= Y83W
XX
XX  JP06253836-A.
XX
XX  13-SEP-1994.
XX
XX  04-MAR-1993; 93JP-0069303.
XX
XX  04-MAR-1993; 93JP-0069303.
XX
XX  (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
XX
XX  WPI: 1994-328987/41.
XX  N-PSDB: AAQ77666.
XX
XX  Variant carbohydrate hydrolase(s) with increased activity -
XX  consists of e.g. alpha-amylase with tyrosine residue in enzyme
XX  centre, useful for mass-prodn. of oligosaccharide(s)
XX
XX  Example 1; Page 15-17; 27pp; Japanese.
XX
XX  AAR63184-7 show the amino acid sequences of variant alpha amylases,
XX  composed by substituting bases 247-249 of the structural gene (see
XX  AAQ77665-8) region, with TTC, TGC, CTC or AAC. These substitutions
XX  result in the 83rd amino acid residue (tyrosine) of the wild type
XX  sequence being changed to phenylalanine, tryptophan, leucine or
XX  asparagine respectively. The substituted amino acid is present in
XX  the active site of the enzyme and confers increased activity on the
XX  enzyme. The variants are useful for the mass production of
XX  oligosaccharides. (see AAR63188 for the variant structure of a
XX  cyclomaltohextrin glucanotransferase).
XX
XX  Sequence 468 AA;
XX
XX  Query Match 67.9%; Score 53; DB 15; Length 468;
XX  Best Local Similarity 53.8%; Pred. No. 2.7;
XX  Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
XX
OY  1 AYHGWMODIYSL 13
XX  |||||:|::||:
XX  79 ayhgwmmknlyki 91
XX
XX  Db
XX
XX  RESULT 13
XX  AAR63186
XX  ID AAR63186 standard; Protein; 452 AA.
XX
XX  AAR63186;
XX
XX  16-JUN-1995 (first entry)
XX
XX  Variant alpha amylase deriv. from Saccharomycopsis fibuligera.
XX
XX  alpha amylase; carbohydrate hydrolase; increased activity;
XX  tyrosine residue; enzyme centre; mass production; oligosaccharide;
XX  variant; cyclomaltohextrin glucanotransferase.
XX
XX  Saccharomycopsis fibuligera.
XX

```

```

XX
XX  Key Location/Qualifiers
FH  MISC-difference 83
FT  /label= Y83L
XX
XX  JP06253836-A.
XX
XX  13-SEP-1994.
XX
XX  04-MAR-1993; 93JP-0069303.
XX
XX  04-MAR-1993; 93JP-0069303.
XX
XX  (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
XX
XX  WPI: 1994-328987/41.
XX  N-PSDB: AAQ77667.
XX
XX  Variant carbohydrate hydrolase(s) with increased activity -
XX  consists of e.g. alpha-amylase with tyrosine residue in enzyme
XX  centre, useful for mass-prodn. of oligosaccharide(s)
XX
XX  Example 1; Page 18-20; 27pp; Japanese.
XX
XX  AAR63184-7 show the amino acid sequences of variant alpha amylases,
XX  composed by substituting bases 247-249 of the structural gene (see
XX  AAQ77665-8) region, with TTC, TGC, CTC or AAC. These substitutions
XX  result in the 83rd amino acid residue (tyrosine) of the wild type
XX  sequence being changed to phenylalanine, tryptophan, leucine or
XX  asparagine respectively. The substituted amino acid is present in
XX  the active site of the enzyme and confers increased activity on the
XX  enzyme. The variants are useful for the mass production of
XX  oligosaccharides. (see AAR63188 for the variant structure of a
XX  cyclomaltohextrin glucanotransferase).
XX
XX  Sequence 452 AA;
XX
XX  Query Match 64.1%; Score 50; DB 15; Length 452;
XX  Best Local Similarity 53.8%; Pred. No. 7.6;
XX  Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
XX
OY  1 AYHGWMODIYSL 13
XX  |||||:|::||:
XX  79 ayhgwmmknlyki 91
XX
XX  Db
XX
XX  RESULT 14
XX  AAP81161
XX  ID AAP81161 standard; Protein; 511 AA.
XX
XX  AAP81161;
XX
XX  12-OCT-1990 (first entry)
XX
XX  Recombinant alpha-amylase.
XX
XX  Alpha amylase; yeast; brewing; baking.
XX
XX  Schwanniomycos castellii ATCC 26076.
XX
XX  EP257115-A.
XX
XX  02-MAR-1988.
XX
XX  21-AUG-1986; 86EP-0111586.
XX
XX  21-AUG-1986; 86EP-0111586.
XX
XX  (HETB ) HEINCKEN TECHNISCH BEHEER BV.
XX
XX  Strasser A, Martens FB, Dohmen J, Hollenberg CP;
XX

```

DR WPI: 1988-057532/09.
 DR N-PSDB; AAN81477.
 XX
 XX Prodn. of amylolytic enzymes by transformed microorganisms -
 PT comprises use of recombinant DNA technology to give transformant
 PT having DNA from donor yeast.
 XX
 PS Disclosure: ; p; English.
 XX
 CC The sequence encoding the protein can be used to construct an
 CC expression vector for the prodn. of alpha amylase in yeast cells.
 CC The transformation and recombinant enzyme produced can be used in
 CC fermentation processes e.g in baking and brewing. The transformed
 CC yeasts are esp. suitable for prodn. of low carbohydrate beers.
 XX
 SQ Sequence 511 AA;

Query Match 64.1%; Score 50; DB 9; Length 511;
 Best Local Similarity 61.5%; Pred. No. 8.7;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 OY 1 AYHGWMQDIYSL 13
 |||||:::|
 Db 114 ayhgywmknidel 126

RESULT 15
 AAP81180
 ID AAP81180 standard; protein; 511 AA.
 XX
 AC AAP81180;
 XX
 DT 12-DEC-1990 (first entry)
 XX
 DE Sequence of alpha-amylase encoded by genomic DNA of Schwanniomycetes
 DE castellii.
 XX
 KW Brewing; beer; breadmaking; biomass.
 XX
 OS Schwanniomycetes castellii.
 XX
 PN EP260404-A.
 PD 23-MAR-1988.
 XX
 PF 17-JUL-1987; 87EP-0110370.
 XX
 PR 17-JUL-1987; 87EP-0110370.
 XX
 PA (HETB) HEINEKEN TEC BEHEER NV.
 XX
 PI Strasser A, Martens FB, Dohmen J, Hollenberg CP;
 DR WPI: 1988-078794/12.
 DR N-PSDB; AAN81525.
 XX
 PT Polypeptide prodn. from cells transformed with yeast DNA -
 PT esp. coding alpha amylase or glucosylase, able to convert
 PT starch to ethanol
 XX
 PS Claim 18; Fig 2A-2C; 74pp; English.
 XX
 CC A cosmid library is constructed from genomic DNA of ATCC26076 and this
 CC screened for inserts contg. the Aa gene by ability to transform
 CC S.cerevisiae GRF 18 to histidine prototrophy. One positive cosmid
 CC (32kb; pYcl) was digested with EcoRI, religated and used to transform
 CC E.coli JA221. Positive transformants contain pYcl-alpha, which includes
 CC a 5kb EcoRI fragment contg. the gene in AAN81525.
 XX
 SQ Sequence 511 AA;

Query Match 64.1%; Score 50; DB 9; Length 511;
 Best Local Similarity 61.5%; Pred. No. 8.7;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 OY 1 AYHGWMQDIYSL 13
 |||||:::|
 Db 114 ayhgywmknidel 126

Search completed: June 13, 2002, 08:53:25
 Job time: 275 sec



•
•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:54:31 ; Search time 57.95 Seconds
(without alignments)
3.793 Million cell updates/sec

Title: US-09-710-339-2_COPY_280_288

Perfect score: 44

Sequence: 1 NAKKSTSGS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgcn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgcn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgcn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgcn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgcn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgcn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	478	1	US-08-720-899-7
2	44	100.0	478	1	US-08-459-610-7
3	44	100.0	478	2	US-08-343-804-7
4	44	100.0	478	2	US-08-339-715A-2
5	44	100.0	478	2	US-08-600-908A-10
6	44	100.0	478	3	US-08-683-838A-10
7	44	100.0	478	4	US-09-182-859-7
8	31	70.5	367	3	US-09-126-741-3
9	31	70.5	490	3	US-09-226-741-1
10	30	68.2	606	4	US-09-041-236-4
11	29	65.9	161	1	US-08-145-995A-13
12	29	65.9	161	2	US-08-451-747-13
13	29	65.9	161	3	US-09-134-852-13
14	29	65.9	870	2	US-09-010-928B-2
15	29	65.9	2089	1	US-08-418-893D-23
16	29	65.9	2089	1	US-08-418-893D-24
17	28	63.6	126	2	US-08-482-728A-10
18	28	63.6	127	4	US-09-160-246-16
19	28	63.6	208	1	US-08-142-897-7
20	28	63.6	248	3	US-08-750-145A-20
21	28	63.6	248	3	US-08-975-698A-24
22	28	63.6	248	4	US-09-417-090-24
23	28	63.6	248	4	US-09-727-578-24
24	28	63.6	280	4	US-09-160-246-14
25	28	63.6	314	1	US-07-704-288C-7
26	28	63.6	314	1	US-08-379-259-7
27	28	63.6	377	1	US-08-525-697-2

28	28	63.6	466	4	US-08-868-435-2	Sequence 2, Appli
29	28	63.6	466	4	US-08-744-231-2	Sequence 2, Appli
30	28	63.6	597	4	US-08-844-274-12	Sequence 12, Appli
31	28	63.6	1188	1	US-08-201-697-4	Sequence 4, Appli
32	28	63.6	1437	3	US-09-061-400-2	Sequence 2, Appli
33	28	63.6	1437	3	US-09-001-273-2	Sequence 2, Appli
34	28	63.6	1453	4	US-08-843-459A-2	Sequence 2, Appli
35	28	63.6	1912	4	US-08-913-832A-2	Sequence 2, Appli
36	27	61.4	152	2	US-08-722-050-12	Sequence 12, Appli
37	27	61.4	154	4	US-08-679-493A-211	Sequence 12, Appli
38	27	61.4	185	2	US-08-691-814B-12	Sequence 12, Appli
39	27	61.4	245	1	US-08-317-707-2	Sequence 2, Appli
40	27	61.4	245	2	US-08-514-921-2	Sequence 2, Appli
41	27	61.4	245	2	US-08-514-921-4	Sequence 4, Appli
42	27	61.4	245	2	US-08-514-921-6	Sequence 6, Appli
43	27	61.4	245	2	US-08-514-921-8	Sequence 8, Appli
44	27	61.4	353	1	US-08-073-807A-17	Sequence 17, Appli
45	27	61.4	380	1	US-08-073-807A-18	Sequence 18, Appli

ALIGNMENTS

RESULT 1
US-08-720-899-7
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57534600 No. 5753460disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowmey Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-720-899-7

Query Match 100.0%; Score 44; DB 1; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Melser, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335.009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 100.0%; Score 44; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
Db 260 NAFKSTSGS 268

RESULT 5
US-08-600-908A-10
Sequence 10, Application US/08600908A
Patent No. 5989169
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 59891690 No. 5989169disk of No. 5989169th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,908A
FILING DATE: 13-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-600-908A-10

Query Match 100.0%; Score 44; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
Db 260 NAFKSTSGS 268

RESULT 6
US-08-683-838A-10
Sequence 10, Application US/08683838A
Patent No. 6022724
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: '-Amylase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 60227240 No. 6022724disk of No. 6022724th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,838A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/600,908
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-683-838A-10

Query Match 100.0%; Score 44; DB 3; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
Db 260 NAFKSTSGS 268

RESULT 7
US-09-182-859-7
Sequence 7, Application US/09182859
Patent No. 6143708
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Borchert, Torben
APPLICANT: Bisgaard-Frantzen, Henrik

```
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/09/182,859
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 0515/96
; EARLIER FILING DATE: 1996-04-30
; EARLIER APPLICATION NUMBER: 0712/96
; EARLIER FILING DATE: 1996-06-28
; EARLIER APPLICATION NUMBER: 0775/96
; EARLIER FILING DATE: 1996-07-11
; EARLIER APPLICATION NUMBER: 1263/96
; EARLIER FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-09-182-859-7
```

```
Query Match          100.0%; Score 44; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 NAFKSTSGS 9
    |||||
Db 260 NAFKSTSGS 268
```

```
RESULT 8
US-09-226-741-3
; Sequence 3, Application US/09226741
; Patent No. 6100037
; GENERAL INFORMATION:
; APPLICANT: Phillips, Stephen C.
; APPLICANT: Lanfear, Jerry
; APPLICANT: Fawcett, Lindsay
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: HUMAN CYCLIC NUCLEOTIDE PDDES
; FILE REFERENCE: PF-0623 US
; CURRENT APPLICATION NUMBER: US/09/226,741
; CURRENT FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 367
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; FEATURE:
; OTHER INFORMATION: HSPDEL0A2
US-09-226-741-3
```

```
Query Match          70.5%; Score 31; DB 3; Length 367;
Best Local Similarity 66.7%; Pred. No. 83;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 NAFKSTSGS 9
    |||||
Db 269 NAFKSTSGS 277
```

```
RESULT 9
US-09-226-741-1
; Sequence 1, Application US/09226741
; Patent No. 6100037
; GENERAL INFORMATION:
; APPLICANT: Phillips, Stephen C.
; APPLICANT: Lanfear, Jerry
; APPLICANT: Fawcett, Lindsay
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: HUMAN CYCLIC NUCLEOTIDE PDDES
```

```
; FILE REFERENCE: PF-0623 US
; CURRENT APPLICATION NUMBER: US/09/226,741
; CURRENT FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 490
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; FEATURE:
; OTHER INFORMATION: HSPDEL0A1
US-09-226-741-1
```

```
Query Match          70.5%; Score 31; DB 3; Length 490;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 NAFKSTSGS 9
    |||||
Db 269 NAFKSTSGS 277
```

```
RESULT 10
US-09-041-236-4
; Sequence 4, Application US/09041236
; Patent No. 6225285
; GENERAL INFORMATION:
; APPLICANT: Luo, Yaling
; APPLICANT: Xiomel, Xu
; TITLE OF INVENTION: Semaphorin K1
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,236
; FILING DATE: March 11, 1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: EXEL98-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-041-236-4
```

```
Query Match          68.2%; Score 30; DB 4; Length 606;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 NAFKSTSGS 9
    |||||
Db 167 NAFKSTSGS 175
```

```
RESULT 11
US-08-145-995A-13
; Sequence 13, Application US/08145995A
; Patent No. 5482850
; GENERAL INFORMATION:
; APPLICANT: CARLOW, CLOTTILDE K.S.
; APPLICANT: PAGE, ANTONY
; TITLE OF INVENTION: METHOD FOR IDENTIFYING ANTI-PARASITIC
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,995A
; FILING DATE: 29-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, DAVID S.
; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 43406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-145-995A-13

Query Match      65.9%; Score 29; DB 1; Length 161;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
Db 99 NAGKNTNGS 107

RESULT 12
US-08-451-747-13
; Sequence 13, Application US/08451747
; Patent No. 5821107
; GENERAL INFORMATION:
; APPLICANT: CARLOW, CLOTTILDE K.S.
; APPLICANT: PAGE, ANTONY
; TITLE OF INVENTION: METHOD FOR IDENTIFYING ANTI-PARASITIC
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: GREGORY D. WILLIAMS; NEW ENGLAND BIOLABS, INC.
; STREET: 32 TOZER ROAD
; CITY: BEVERLY
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 01915
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,747
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/145,995
; FILING DATE: 29-OCT-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-046-DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 927-5054
; TELEFAX: (508) 927-1705
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-451-747-13
```

```
Query Match      65.9%; Score 29; DB 2; Length 161;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
Db 99 NAGKNTNGS 107
```

```
RESULT 13
US-09-134-852-13
; Sequence 13, Application US/09134852
; Patent No. 6127148
; GENERAL INFORMATION:
; APPLICANT: CARLOW, CLOTTILDE K.S.
; APPLICANT: PAGE, ANTONY
; TITLE OF INVENTION: METHOD FOR IDENTIFYING ANTI-PARASITIC
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,852
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/145,995
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, DAVID S.
; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 43406
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (617) 523-3400
TELEFAX: (617) 523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 161 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-134-852-13

Query Match 65.9%; Score 29; DB 3; Length 161;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NAFKSTGS 9
DB 99 NAKNTNGS 107

RESULT 14
US-09-010-928B-2
; Sequence 2, Application US/09010928B
; Patent No. 5994099
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V
; APPLICANT: Hayashi, Cheryl Y
; TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
; TITLE OF INVENTION: CODING THEREFOR
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 GATEHOUSE RD. SUITE 500E
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,928B
; FILING DATE: 22-JAN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M
; REGISTRATION NUMBER: 28977
; REFERENCE/DOCKET NUMBER: 1447-109P
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 870 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-010-928B-2

Query Match 65.9%; Score 29; DB 2; Length 870;
Best Local Similarity 66.7%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NAFKSTGS 9
DB 49 NALMSTGS 57

RESULT 15
US-08-418-893D-23
; Sequence 23, Application US/08418893D

; Patent No. 5559220
; GENERAL INFORMATION:
; APPLICANT: ROESSLER, PAUL G
; APPLICANT: OHLROGGE, JOHN B
; TITLE OF INVENTION: GENE THAT ENCODES ACETYL-COENZYME A
; TITLE OF INVENTION: CARBOXYLASE FROM CYCLOTETRA CRYPTIC
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NATIONAL RENEWABLE ENERGY LABORATORY
; STREET: 1617 Cole Blvd.
; CITY: Golden
; STATE: CO
; COUNTRY: USA
; ZIP: 80401-3393
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,893D
; FILING DATE: April 7, 1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/104,938
; FILING DATE: September 14, 1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: O'CONNOR, EDNA
; REGISTRATION NUMBER: 29,252
; REFERENCE/DOCKET NUMBER: MRI/NREL IR 92-48CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303-231-1000
; TELEFAX: 303-231-1098
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2089 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOHETICAL: NO
; ANTI-SENSE: NO
; US-08-418-893D-23

Query Match 65.9%; Score 29; DB 1; Length 2089;
Best Local Similarity 85.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 FKSTGS 9
DB 502 FKPTGS 508

Search completed: June 13, 2002, 08:54:32
Job time: 317 sec

Fri Jun 14 10:36:46 2002

us-09-710-339-2_copy_280_288.ra1

Page 7

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:12 ; Search time 125.81 Seconds
(without alignments)
12.375 Million cell updates/sec

Title: US-09-710-339-2_COPY_280_288

Sequence: 1 MAFKSTSGS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriapi:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44	100.0	499	3	Q96TH4	Q96th4 aspergillus
2	40	90.9	623	3	Q9UV09	Q9uv09 emericella
3	37	84.1	2276	2	Q93TV6	Q93tv6 staphylococ
4	34	77.3	102	3	P87271	P87271 saccharomyc
5	34	77.3	902	6	Q9BDP9	Q9bdt9 echinops te
6	34	77.3	915	6	Q9BDM8	Q9bdt8 tenrec ecan
7	34	77.3	1146	6	Q46488	Q46488 alouatta se
8	33	75.0	297	16	Q9KBL4	Q9kbl4 bacillus ha
9	33	75.0	490	3	Q9UV07	Q9uv07 emericella
10	33	75.0	631	3	Q92394	Q92394 cryptococcu
11	33	75.0	688	5	Q917J5	Q917j5 drosophila
12	33	75.0	690	5	Q9NH72	Q9nh72 drosophila
13	33	75.0	690	5	Q9VFT4	Q9vft4 drosophila
14	32	72.7	100	17	Q9VAT6	Q9vat6 aeropyrum p
15	32	72.7	196	11	Q9EOM8	Q9eom8 mus musculu
16	32	72.7	220	15	Q900X2	Q90qx2 human immun

17	32	72.7	399	13	Q98SE8	Q98se8 gallus gall
18	32	72.7	427	17	Q97C83	Q97c83 thermoplasma
19	32	72.7	513	10	Q81992	Q81992 hordeum vul
20	32	72.7	513	10	Q40029	Q40029 hordeum vul
21	32	72.7	681	5	Q20353	Q20353 caenorhabdi
22	32	72.7	760	3	Q9P3E0	Q9p3e0 neurospora
23	32	72.7	937	4	Q9H3Q4	Q9h3q4 homo sapien
24	32	72.7	1198	5	Q9U5Y2	Q9u5y2 drosophila
25	32	72.7	1200	5	Q9V8U8	Q9v8u8 drosophila
26	32	72.7	1200	5	Q9U5J1	Q9u5j1 drosophila
27	32	72.7	1200	5	Q9U472	Q9u472 drosophila
28	32	72.7	1638	4	Q9Y6U4	Q9y6u4 homo sapien
29	32	72.7	1649	16	Q9CF22	Q9cf22 lactococcus
30	32	72.7	1712	4	Q9Y6U5	Q9y6u5 homo sapien
31	31	70.5	194	10	Q93ZG2	Q93zg2 arabidopsis
32	31	70.5	221	8	Q79611	Q79611 gomphus flo
33	31	70.5	234	16	Q9ZE22	Q9ze22 listeria in
34	31	70.5	419	13	Q92043	Q92043 crotilus at
35	31	70.5	471	10	Q9XIB2	Q9xib2 arabidopsis
36	31	70.5	480	4	Q9H0D4	Q9h0d4 homo sapien
37	31	70.5	490	4	Q9NY45	Q9ny45 homo sapien
38	31	70.5	492	4	Q9NSM4	Q9nsn4 homo sapien
39	31	70.5	492	16	Q9CE95	Q9ce95 lactococcus
40	31	70.5	562	11	Q9WVG4	Q9wvg4 rattus norv
41	31	70.5	572	5	Q26042	Q26042 penaeus jap
42	31	70.5	576	4	Q9HB46	Q9hb46 homo sapien
43	31	70.5	595	10	Q9AR86	Q9ar86 populus x c
44	31	70.5	607	5	Q9U5H3	Q9u5h3 caenorhabdi
45	31	70.5	621	11	Q61866	Q61866 mus musculu

ALIGNMENTS

RESULT 1
ID Q96TH4 PRELIMINARY; PRT; 499 AA.
AC Q96TH4;
DT 01-DEC-2001 (TRENBLER, 19, Created)
DT 01-DEC-2001 (TRENBLER, 19, Last sequence update)
DT 01-DEC-2001 (TRENBLER, 19, Last annotation update)
DE TAKA-AMYLASE A (EC 3.2.1.1) (ALPHA-AMYLASE).
GN AMYA OR AMY1.
OS Aspergillus oryzae, and
OS Aspergillus oryzae, and
OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062, 5059;
RN [1]
RP SEQUENCE FROM N.A. STRAIN-RIB40;
RC SPECIES=A.oryzae; STRAIN=RIB40;
RX MEDLINE=20289310; PubMed=10830498;
RA Gomi K., Akeno T., Minetoki T., Ozeki K., Kumagai C., Okazaki N., Iimura Y.;
RT "Molecular cloning and characterization of a transcriptional activator gene, amyR, involved in the amylolytic gene expression in Aspergillus oryzae.";
RT Biosci. Biotechnol. Biochem. 64:816-827(2000).
RL [2]
RN SEQUENCE FROM N.A.
RC SPECIES=A. flavus; STRAIN=86-10D;
RA Fakhouy A.M., Woloshuk C.P.;
RT "Amyl, the alpha-amylase gene of Aspergillus flavus: Involvement in allicoxin biosynthesis in maize kernels.";
RL Phytopathology 89:908-914(1999).
DR EMBL; AB021876; BAB95703.1; -
SO SOURCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;

Query Match 100.0%; Score 44; DB 3; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
 |||||
 DB 281 NAFKSTSGS 289

RESULT 2
 Q9UV09 PRELIMINARY; PRT; 623 AA.
 AC Q9UV09;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE ALPHA-AMYLASE.
 GN AMYB.
 OS *Emmericella nidulans* (*Aspergillus nidulans*).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutriales; Trichocomaceae; *Emmericella*.
 OX NCBI_TaxID=5072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Boase N.A., Murphy R.L., Kelly J.M.;
 RA "Characterisation and regulation of an alpha-amylase gene in
 RT *Aspergillus nidulans*.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF208224; AAF17100.1; -
 DR HSSP: p56271; 2AAA
 DR InterPro: IPR000461; Alpha_amylase.
 DR InterPro: IPR002044; CBD_4.
 DR Pfam: PF00128; alpha-amylase; 1.
 DR Pfam: PF00686; CBD_4; 1.
 DR PRINTS: PR00110; ALPHAMYLASE.
 DR ProDom: PD001568; CBD_4; 1.
 RN SEQUENCE 623 AA; 68598 MW; 822F002C37F5A9A4 CRC64;

Query Match 90.9%; Score 40; DB 3; Length 623;
 Best Local Similarity 88.9%; Pred. No. 4.1;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NAFKSTSGS 9
 |||||
 DB 279 NAFKSTSGS 287

RESULT 3
 Q93TY6 PRELIMINARY; PRT; 2276 AA.
 AC Q93TY6;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE BIOFILM-ASSOCIATED SURFACE PROTEIN.
 GN BAP.
 OS *Staphylococcus aureus*.
 OC Bacteria; Firmicutes; *Bacillus*/Clostridium group;
 OC *Bacillus*/Staphylococcus group; *Staphylococcus*.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V329;
 RX MEDLINE=21189316; PubMed=11292810;
 RA Cucarella C., Solano C., Valle J., Amorena B., Lasa I.,
 RA Penades J.R.;
 RA "Bap, a *Staphylococcus aureus* Surface Protein Involved in Biofilm
 RT Formation.";
 RL J. Bacteriol. 183:2888-2896(2001).
 DR EMBL: AF288402; AAK38834.1; -
 SO SEQUENCE 2276 AA; 238792 MW; 8623C60472CBOC0F CRC64;

Query Match 84.1%; Score 37; DB 2; Length 2276;
 Best Local Similarity 77.8%; Pred. No. 65;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NAFKSTSGS 9
 |||||
 DB 548 NAFKSTSGS 556

RESULT 4
 P87271 PRELIMINARY; PRT; 102 AA.
 AC P87271;
 DT 01-JUL-1997 (TREMblrel. 04, Created)
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE YDR455C.
 GN YDR455C.
 OS *Saccharomyces cerevisiae* (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; *Saccharomyces*.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dietrich F.S.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
 RA Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
 RA Hunnicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosesdale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Stroff N.,
 RA Winant A., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Jia Y., Cherry J.M.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U33007; AAB6489.1; -
 DR SGD: S0002863; YDR455C.
 SO SEQUENCE 102 AA; 10691 MW; 0EFC38552A82BB9 CRC64;

Query Match 77.3%; Score 34; DB 3; Length 102;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 FKSTSGS 9
 |||||
 DB 15 FKSTSGS 21

RESULT 5
 Q9BDT9 PRELIMINARY; PRT; 902 AA.
 AC Q9BDT9;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
 DE BRCA1 (FRAGMENT)
 OS *Echinops telfairi* (Lesser hedgehog tenrec).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Insectivora; Tenrecidae; Tenrecinae; *Echinops*.
 OX NCBI_TaxID=9371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21082081; PubMed=11214318;
 RA Maden O., Scally M., Donady C., Kao D., DeBry R., Adkins R.,
 RA Amurte H., Stanhope M., de Jong W., Springer M.;
 RT "Parallel adaptive radiations in two major clades of placental
 RT mammals.";
 RL Nature 409:610-614(2001).
 DR EMBL: AF284025; AAK15612.1; -
 DR InterPro: IPR002378; Brst_cancerI.

DR InterPro: IPR001993; Mitoch_carrier.
 DR PRINTS: PR00493; BRSTCANCER1.
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
 FT NON_TER 1
 FT NON_TER 902 902
 SQ SEQUENCE 902 AA: 98813 MW: CD8E27EA1A3392CD CRC64;

Query Match 77.3%; Score 34; DB 6; Length 902;
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 |||||
 Db 738 NAFKGTSSS 746

RESULT 6
 Q9BDT8 PRELIMINARY; PRT; 915 AA.
 AC Q9BDT8;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE BRCA1 (FRAGMENT).
 OS Tenrec ecaudatus (tailless tenrec).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Insectivora; Tenrecidae; Tenrecinae; Tenrec.
 OX NCBI_TaxID=94439;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21082081; PubMed=11214318;
 RA Maden O., Scally M., Donady C., Kao D., Debry R., Adkins R.,
 RA Amrine H., Stanhope M., de Jong W., Springer M.;
 RT "Parallel adaptive radiations in two major clades of placental
 mammals."
 RL Nature 409:610-614(2001).
 RT EMBL: AF284026; AAK15613.1; -
 DR InterPro: IPR002378; Brst_cancert.
 DR PRINTS: PR00493; BRSTCANCER1.
 FT NON_TER 1
 FT NON_TER 915 915
 SQ SEQUENCE 915 AA: 100159 MW: A9D62AA3E83E8F36 CRC64;

Query Match 77.3%; Score 34; DB 6; Length 915;
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 |||||
 Db 743 NAFKGTSSS 751

RESULT 7
 Q46488 PRELIMINARY; PRT; 1146 AA.
 AC Q46488;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE BREAST CANCER TYPE 1 SUSCEPTIBILITY PROTEIN (BREAST AND OVARIAN CANCER
 DE SUSCEPTIBILITY PROTEIN) (FRAGMENT).
 GN BRCA1.
 OS Alouatta seniculus (Howler monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Alouattinae;
 OC Alouatta.
 OX NCBI_TaxID=9503;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98122577; PubMed=9462745;
 RA Hacia J.G., Makalowski W., Edgemon K., Erdos M.R., Robbins C.M.,

RA Fodor S.P.A., Brody L.C., Collins F.S.;
 RT "Evolutionary sequence comparisons using high-density oligonucleotide
 RT arrays."
 RL Nat. Genet. 18:155-158(1998).
 CC -I- FUNCTION: NOT KNOWN, MAY REGULATE GENE EXPRESSION. MAY PLAY A ROLE
 CC IN CELL PROLIFERATION AND DIFFERENTIATION (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 DR EMBL: AF019079; AAC39587.1; -
 DR InterPro: IPR002378; Brst_cancert.
 DR PRINTS: PR00493; BRSTCANCER1.
 KW DNA-binding; Nuclear protein; Anti-oncogene.
 FT NON_TER 1
 FT DOMAIN 278 285 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 200 203 POLY-SER.
 FT DOMAIN 815 818 POLY-SER.
 FT DOMAIN 1001 1004 POLY-GLU.
 FT DOMAIN 996 1005 LEUCINE-ZIPPER (POTENTIAL).
 FT NON_TER 1146 1146
 SQ SEQUENCE 1146 AA: 128119 MW: BFD07EBBF851890E CRC64;

Query Match 77.3%; Score 34; DB 6; Length 1146;
 Best Local Similarity 77.8%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 |||||
 Db 809 NAFKGTSSS 817

RESULT 8
 Q9KBL4 PRELIMINARY; PRT; 297 AA.
 AC Q9KBL4;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE TRANSMEMBRANE LIPOPROTEIN.
 DE LIPIC OR BH1912.
 GN Bacillus halodurans.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horiuchi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001513; BAB05631.1; -
 DR InterPro: IPR000515; BPD_transp.
 DR Pfam: PF00528; BPD_transp.1.
 DR PROSITE: PS00402; BPD_TRANS_PLM_MEMBER; UNKNOWN_1.
 DR PROSITE: PS00402; BPD_TRANS_PLM_MEMBER; UNKNOWN_1.
 KW Lipoprotein; Complete proteome.
 SQ SEQUENCE 297 AA: 33588 MW: BC744C0A5E8FAEE CRC64;

Query Match 75.0%; Score 33; DB 16; Length 297;
 Best Local Similarity 75.0%; Pred. No. 59;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NAFKSTSG 8
 |||||
 Db 244 NAFKGTSG 251

RESULT 9
 Q9UV07 PRELIMINARY; PRT; 490 AA.

```
AC Q9UV07;
DC 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE AMYA.
GN AMYA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RA Roase N.A., Murphy R.M., Kelly J.M.;
RT "An amylase cluster in Aspergillus nidulans."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF208225; AAF17103.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR000461; Alpha-amylase.
DR Pfam; PF00128; alpha-amylase_1.
DR Pfam; PF02806; alpha-amylase_C; 1.
DR PRINTS; PR00110; ALPHAAMYLASE.
SQ SEQUENCE 490 AA; 54249 MW; A891C4ACEAB5305 CRC64;

Query Match          75.0%; Score 33; DB 3; Length 490;
Best Local Similarity 87.5%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AFKSTSGS 9
DB 274 AFSTSGS 281

RESULT 10
ID 092394 PRELIMINARY; PRT; 631 AA.
AC 092394;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ALPHA-AMYLASE PRECURSOR.
OS AMY-CS2.
GN Cryptococcus sp. S-2.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
OC Microsporid Hymenomycetes; Cryptococcus.
OX NCBI_TaxID=87049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S-2;
RX MEDLINE=96433120; PubMed=8836148;
RA Iefuji H., Chino M., Kato M., Iimura Y.;
RT "Raw starch digesting and thermostable alpha-amylase from the yeast
RT Cryptococcus sp. S-2: purification, characterization, cloning and
RT sequencing."
RL Biochem. J. 318:989-996(1996).
DR EMBL; D83540; BAA12011.1; -.
DR EMBL; D83540; BAA12010.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro; IPR000461; Alpha-amylase.
DR InterPro; IPR002044; CBD_4.
DR Pfam; PF00128; alpha-amylase_1.
DR Pfam; PF00686; CBD_4; 1.
DR ProDom; PD001568; CBD_4; 1.
DR Signal.
KW Signal.
FT CHAIN 1 20 ALPHA-AMYLASE.
FT SIGNAL 21 631
SQ SEQUENCE 631 AA; 67658 MW; 819687B6E1D707E5 CRC64;

Query Match          75.0%; Score 33; DB 3; Length 631;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 NAFKSTSGS 9
DB 295 NAFKSTSGS 303

RESULT 11
ID 091705 PRELIMINARY; PRT; 688 AA.
AC 091705;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE RIN PROTEIN.
GN RIN OR CG9412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang O., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazer G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Adayant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor J.C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostali D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhan G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003701; AAG22151.1; -.
DR FlyBase; FBgn0015778; rin.
DR InterPro; IPR002075; NTF2.
DR InterPro; IPR000504; RRM.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PSS0102; RRM; 1.
SQ SEQUENCE 688 AA; 74713 MW; 9E329949E7248C19 CRC64;

Query Match          75.0%; Score 33; DB 5; Length 688;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
```

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
| | | | |
DB 385 NLFKSTSSS 393

RESULT 12
Q9NH72 PRELIMINARY; PRT; 690 AA.

DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RASPUTIN.
GN RIN OR CG9412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20191744; PubMed=10725247;
RA Pazman C., Mayes C.A., Panto M., Haynes S.R., Mlodzik M.;
RT "Rasputin, the Drosophila homolog of the RasGAP SH3 binding protein,
functions in Ras and Rho mediated signaling.";
RL Development 127:1715-1725(2000).
DR EMBL; AF21031; AAF68949.1; -;
DR Flybase; FBgn0015778; rin.
DR InterPro; IPR002075; NTF2.
DR InterPro; IPR005054; RRM.
DR Pfam; PF00076; Rrm; 1.
DR SMART; SMO0360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
SQ SEQUENCE 690 AA; 74985 MW; 40C5BD7E95FA253 CRC64;

Query Match 75.0%; Score 33; DB 5; Length 690;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
| | | | |
DB 387 NLFKSTSSS 395

RESULT 13
Q9VFT4 PRELIMINARY; PRT; 690 AA.

AC Q9VFT4
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CG9412 PROTEIN (LD31194P).
GN RIN OR CG9412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RX Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
Georgie R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champs M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., Meleod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui Y., Reese M.G.,
RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Paclet J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003701; AAF54965.1; -;
DR EMBL; AY058617; AAL18486.1; -;
DR Flybase; FBgn0015778; rin.
DR InterPro; IPR002075; NTF2.
DR InterPro; IPR005054; RRM.
DR Pfam; PF02136; rtm; 1.
DR Pfam; PF00076; Rrm; 1.
DR SMART; SMO0360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
SQ SEQUENCE 690 AA; 74943 MW; 40C5AC6F9FAEB353 CRC64;

Query Match 75.0%; Score 33; DB 5; Length 690;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
| | | | |
DB 387 NLFKSTSSS 395

RESULT 14
Q9YAT6 PRELIMINARY; PRT; 100 AA.

AC Q9YAT6
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE HYPOHETICAL 10.4 KDA PROTEIN APE1858.
GN APE1858.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcates; Desulfurococcaceae;
OC Aeropyrum.
NCBI_TaxID=56636;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1:
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 DR EMBL: AP000062; BAA0862.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 100 AA; 10356 MM; GC17D0ED584681EA CRC64;

Query Match 72.7%; Score 32; DB 17; Length 100;
 Best Local Similarity 77.8%; Pred. No. 32;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 ||| 1:111
 DB 43 NAFCSSTSGS 51

RESULT 15
 O9EQM8 PRELIMINARY; PRT; 196 AA.
 AC O9EQM8;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE NUCLEOPORIN 98 (FRAGMENT).
 GN NUP98.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wu X., van Deursen J.M.A.;
 RT "Alternative splicing of mouse Nup98 mRNA.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF201386; AAC42323.1;
 DR MGD: MGI:109404; Nup98.
 KW Porin.
 FT NON_TER 1
 FT NON_TER 196
 SQ SEQUENCE 196 AA; 19483 MM; 7F508FBD4D5AE3D3 CRC64;

Query Match 72.7%; Score 32; DB 11; Length 196;
 Best Local Similarity 77.8%; Pred. No. 63;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 ||| 1:11111
 DB 125 NPFSTSGS 133

Search completed: June 13, 2002, 09:11:15
 Job time: 1060 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 09:11:53 ; Search time 38.1 Seconds
(without alignments)
9.146 Million cell updates/sec

Title: US-09-710-339-2_COPY_280_288

Perfect score: 44

Sequence: 1 NAFKSTGS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

105224

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	44	100.0	AMYA_ASPAW	Q02905 aspergillus
2	44	100.0	AMYA_ASPOR	P10529 aspergillus
3	44	100.0	AMYE_ASPAW	Q02906 aspergillus
4	44	100.0	AMY_ASPRH	P30292 aspergillus
5	32	72.7	FLG6_BORBU	O51715 borrelia bu
6	32	72.7	SLAF_AERSA	P35823 aeromonas s
7	32	72.7	NU98_HUMAN	P52948 homo sapien
8	31	70.5	ZRF1_MOUSE	P54103 mus musculu
9	31	70.5	CID_DROME	P19538 drosophila
10	30	68.2	TPMT_MOUSE	O55060 mus musculu
11	30	68.2	TPMT_MOUSE	O94X22 mus spreus
12	30	68.2	RTVA_RICCN	Q94187 rickettsia
13	30	68.2	E13C_TOBAC	P23432 nicotiana t
14	30	68.2	DPAL_SALTY	P40817 salmoneilla
15	30	68.2	FABR_ARATH	P52410 arabidopsis
16	30	68.2	AMYA_ASPNG	P56271 aspergillus
17	30	68.2	FLHA_HELPJ	O96758 helicobacte
18	30	68.2	FLHA_HELPJ	O96758 helicobacte
19	30	68.2	APAF_BRARE	O91958 brachydanio
20	29	65.9	RBS1_RHOSH	P27998 rhodobacter
21	29	65.9	VG14_BPPI2	P11188 bacterioph
22	29	65.9	VG14_BPPI2	P07539 bacterioph
23	29	65.9	CYPH_ECHGR	P14088 echinococcu
24	29	65.9	PEI_STRCH	O06118 streptococcu
25	29	65.9	YT22_CAEEL	O10932 caenorhabdi
26	29	65.9	MOVV_AAVLE	P03595 alfalfa mos
27	29	65.9	MOVV_AAVMA	P03672 alfalfa mos
28	29	65.9	MOVV_AAVST	P03596 alfalfa mos
29	29	65.9	MOVV_AAVYS	P24256 alfalfa mos
30	29	65.9	Y306_MYCPN	P73343 mycoplasma
31	29	65.9	ODO2_YEAST	P19262 saccharomyc
32	29	65.9	RK11_SRCEC	O02723 secale cere
33	29	65.9	RRP1_CDVO	P06940 canine dist

34	29	65.9	710	1	IRAL_MOUSE	Q62406 mus musculu
35	29	65.9	759	1	RRP3_IANN	P21428 influenza a
36	29	65.9	759	1	RRP3_IABUD	P26104 influenza a
37	29	65.9	759	1	RRP3_IACHI	P11486 influenza a
38	29	65.9	759	1	RRP3_IAPOM	Q82573 influenza a
39	29	65.9	759	1	RRP3_IAPPR	P12445 influenza a
40	29	65.9	759	1	RRP3_IAPPW	P18883 influenza a
41	29	65.9	759	1	RRP3_IAGU2	P26115 influenza a
42	29	65.9	759	1	RRP3_IAGU1	P26116 influenza a
43	29	65.9	759	1	RRP3_IAMK6	P26105 influenza a
44	29	65.9	759	1	RRP3_IAMKO	P26106 influenza a
45	29	65.9	759	1	RRP3_IAMKE	P16884 influenza a

ALIGNMENTS

RESULT	ID	AMYA_ASPAW	STANDARD	PRT	498 AA.
AC	Q02905				
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	01-OCT-1996	(Rel. 34, Last annotation update)			
DE	Alpha-amylase A precursor (EC 3.2.1.1) (1.4-alpha-D-glucan glucanohydrolase A).				
DE	glucanohydrolase A).				
OS	AMYA.				
OS	Aspergillus awamori.				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
OC	Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				
OX	NCBI_TaxID=105351;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=UVR143P;				
RX	MEDLINE=90254827; PubMed=2340591;				
RA	Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,				
RA	Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,				
RA	Berka R.M.;				
RT	"Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus niger var. awamori.";				
RL	Curr. Genet. 17:203-212(1990).				
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides.				
CC	-1- CORFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: X52755; CAA36966.1; -				
DR	HSSP: P10529; 7TAA.				
DR	InterPro: IPR000461; Alpha-amylase.				
DR	Pfam: PF00128; alpha-amylase; 1.				
KW	Glycoprotein; Signal; Multigene family.				
KW	GLYCOPROTEIN; Signal; Multigene family.				
FT	SIGNAL	1	21	POTENTIAL.	
FT	CHAIN	22	498	ALPHA-AMYLASE A.	
FT	ACT_SITE	227	227	BY SIMILARITY.	
FT	ACT_SITE	231	231	BY SIMILARITY.	
FT	ACT_SITE	251	251	BY SIMILARITY.	
FT	ACT_SITE	318	318	BY SIMILARITY.	
FT	DISULFID	51	59	BY SIMILARITY.	
FT	DISULFID	171	185	BY SIMILARITY.	
FT	DISULFID	261	304	BY SIMILARITY.	
FT	DISULFID	461	496	BY SIMILARITY.	
FT	CARBOHYD	218	218	N-LINKED (GLCNAC...) (POTENTIAL).	
SO	SEQUENCE	498 AA;	54880 MW;	765851 IBC01A8A01 CRC64;	

Query Match 100.0%; Score 44; DB 1; Length 498;
 Best Local Similarity 100.0%; Prod. No. 0 088;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 |||||
 DB 281 NAFKSTSGS 289

RESULT 2
 ID AMYA_ASFOR STANDARD: PRT: 499 AA.
 AC P10529; P11763; Q00250;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
 DE (1,4-alpha-D-glucan glucanohydrolase).
 GN AMY1 AND AMY2 AND AMY3.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 63303;
 RA MEDLINE=89237897; PubMed=2785629;
 RA Wiesel S., Lachmund A., Wildhardt G., Rutkowski E.;
 RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
 RT Intron-exon organization.";
 RL Mol. Microbiol. 3:3-14(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89378767; PubMed=2789162;
 RA Genes M.J., Dove M.J., Seligy V.L.;
 RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
 RT containing eight introns.";
 RL Gene 79:107-117(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tada S., Iimura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
 RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
 RT of Aspergillus oryzae.";
 RL Agric. Biol. Chem. 53:593-599(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90128276; PubMed=2612911;
 RA Tsukagoshi N., Furukawa M., Nagaba H., Kirta N., Tsuboi A.,
 RA Ueda S.;
 RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
 RT evidence for multiple related genes.";
 RL Gene 84:319-327(1989).
 RN [5]
 RP SEQUENCE OF 22-499.
 RA Toda H., Kondo K., Narita K.;
 RT "The complete amino acid sequence of Taka-amylase A.";
 RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
 RN [6]
 RP SEQUENCE OF 206-225.
 RX MEDLINE=74001521; PubMed=4733850;
 RA Isemura S., Ikenaka T.;
 RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
 RT A with trypsin and chymotrypsin.";
 RL J. Biochem. 74:1-10(1973).
 RN [7]
 RP SEQUENCE OF 433-499.
 RA Narita K.;
 RT Proc. Jpn. Acad., B, Phys. Biol. Sci. 51:285-290(1975).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=80227691; PubMed=6156152;

RA Matsura Y., Kusunoki M., Harada W., Tanaka N., Iga Y., Yasuoka N.,
 RA Toda H., Narita K., Kakudo M.;
 RT "Molecular structure of taka-amylase A. I. Backbone chain folding at
 RT 3-A resolution.";
 RL J. Biochem. 87:1555-1558(1980).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RX MEDLINE=84213370; PubMed=6609921;
 RA Matsura Y., Kusunoki M., Harada W., Kakudo M.;
 RT "Structure and possible catalytic residues of Taka-amylase A.";
 RL J. Biochem. 95:697-702(1984).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).
 RX MEDLINE=97428212; PubMed=9283074;
 RA Brzozowski A.M., Davies G.J.;
 RT "Structure of the Aspergillus oryzae alpha-amylase complexed with the
 RT inhibitor acarbose at 2.0-A resolution.";
 RL Biochemistry 36:10837-10845(1997).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1- MISCELLANEOUS: THE SEQUENCE OF AMY1 AND AMY2 IS SHOWN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X12725; CAA31218.1; -;
 DR EMBL: X12726; CAA31219.1; -;
 DR EMBL: X12727; CAA31220.1; -;
 DR EMBL: D00434; BAA00336.1; -;
 DR EMBL: M33218; AAA32708.1; -;
 DR PIR: S04548; ALAS1.
 DR PIR: S04549; ALAS3.
 DR PIR: A10627; A10627.
 DR PIR: A10627; A10627.
 DR PIR: JK0201; JK0201.
 DR PIR: JS0240; JS0240.
 DR PIR: JT0466; JT0466.
 DR PDB: 2TAA; 15-OCT-89.
 DR PDB: 6TAA; 31-OCT-93.
 DR PDB: 7TAA; 25-NOV-98.
 DR GlycoSiteDB: P10529;
 DR InterPro: IPR000461; Alpha-amylase.
 DR Pfam: PF00128; alpha-amylase, 1.
 KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
 KW Glycoprotein; Signal; Multigene family; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 499
 FT ACT_SITE 227 227
 FT ACT_SITE 231 231
 FT ACT_SITE 251 251
 FT ACT_SITE 318 318
 FT DISULFID 51 39
 FT DISULFID 171 185
 FT DISULFID 261 304
 FT DISULFID 461 496
 FT CARBOHYD 218 218
 FT
 FT VARIANT 56 56
 FT VARIANT 172 172
 FT VARIANT 93 94
 FT CONFLICT 106 106
 FT CONFLICT 184 184
 FT CONFLICT 195 195
 FT CONFLICT 255 255
 FT CONFLICT 291 291
 FT CONFLICT 345 345
 N-LINKED (GLCNAC. . .).
 /FTid=CAR.000125.
 Q -> R (IN AMY3).
 F -> L (IN AMY3).
 TT -> DC (IN REF. 5).
 Q -> T (IN REF. 5).
 D -> Y (IN REF. 3).
 P -> L (IN REF. 3).
 G -> V (IN REF. 3).
 D -> H (IN REF. 4).
 I -> L (IN REF. 5).

```

FT CONFLICT 370 370 L -> A (IN REF. 4).
FT CONFLICT 406 409 WEIY -> PYI (IN REF. 5).
FT CONFLICT 448 448 G -> S (IN REF. 5).
FT CONFLICT 497 497 S -> SD (IN REF. 5 AND 7).
FT HELIX 24 27
FT TURN 28 29
FT STRAND 32 35
FT HELIX 37 40
FT STRAND 42 42
FT TURN 43 44
FT HELIX 53 55
FT STRAND 61 61
FT HELIX 63 67
FT TURN 68 69
FT HELIX 70 74
FT TURN 75 77
FT STRAND 80 83
FT STRAND 87 89
FT STRAND 94 94
FT TURN 95 96
FT STRAND 97 97
FT TURN 100 101
FT STRAND 105 111
FT TURN 113 115
FT TURN 118 119
FT HELIX 120 130
FT TURN 131 132
FT STRAND 134 139
FT STRAND 143 143
FT STRAND 146 146
FT HELIX 150 152
FT TURN 155 156
FT STRAND 159 159
FT HELIX 164 166
FT STRAND 167 167
FT STRAND 172 172
FT HELIX 179 184
FT STRAND 186 188
FT STRAND 193 194
FT STRAND 196 197
FT TURN 199 200
FT HELIX 202 219
FT TURN 220 220
FT STRAND 223 226
FT TURN 227 228
FT HELIX 229 231
FT TURN 234 235
FT HELIX 236 244
FT STRAND 247 250
FT HELIX 257 265
FT TURN 266 266
FT STRAND 269 271
FT HELIX 273 283
FT TURN 286 287
FT HELIX 290 303
FT STRAND 307 309
FT TURN 311 312
FT TURN 317 318
FT HELIX 322 325
FT STRAND 329 341
FT STRAND 345 349
FT TURN 350 351
FT HELIX 352 354
FT TURN 355 355
FT TURN 361 364
FT HELIX 368 371
FT TURN 372 372

```

Query Match
Best Local Similarity 100.0%; Score 44; DB 1; Length 499;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9

```

DB 281 NAFKSTSGS 289
|||||
RESULT 3
ID AMYB_ASPAW STANDARD; PRT; 499 AA.
AC 002906;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alpha-amyrase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase B).
GN AMYB.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UW143F;
RX MEDLINE=90254827; Pubmed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amyase genes
RT from Aspergillus niger var. awamori."
RL Curr. Genet. 17:203-212(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52756; CAA36967.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro: IPR000461; Alpha-amyase.
DR Pfam: PF00128; alpha-amyase.1.
KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 499
FT ACT_SITE 227 227 BY SIMILARITY.
FT ACT_SITE 231 231 BY SIMILARITY.
FT ACT_SITE 251 251 BY SIMILARITY.
FT ACT_SITE 318 318 BY SIMILARITY.
FT DISULFID 51 59 BY SIMILARITY.
FT DISULFID 171 185 BY SIMILARITY.
FT DISULFID 261 304 BY SIMILARITY.
FT DISULFID 461 496 BY SIMILARITY.
FT CARBOHYD 218 218 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 499 AA; 54921 MW; 740B96B11BC01A8A CRC64;

```

Query Match
Best Local Similarity 100.0%; Score 44; DB 1; Length 499;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9

RESULT 4
AMY ASPSH

```

ID  AMY_ASPPH  STANDARD;  PRT;  499 AA.
AC  P30292;
DT  01-APR-1993 (Rel. 25, Created)
DT  01-APR-1993 (Rel. 25, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Alpha-amyrase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE  glucanohydrolase).
GN  AMY.
OS  Aspergillus shirousami.
OC  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC  Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX  NCBI_TaxID=5070;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=9232146; PubMed=1368777;
RA  Shiba Y., Tamura G., Ishikawa T., Hara S.;
RT  "Cloning of the alpha-amyrase cDNA of Aspergillus shirousami and its
RT  expression in Saccharomyces cerevisiae."
RL  Biosci. Biotechnol. Biochem. 56:174-179(1992).
CC  -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC  linkages in oligosaccharides and polysaccharides.
CC  -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC  -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC  KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: D10461; BAA01255.1; -.
DR  PIR: J50663; J50663.
DR  HSSP: P10529; 7TAA.
DR  InterPro: IPR000461; Alpha-amyrase.
DR  Pfam: PF00128; alpha-amyrase; 1.
KM  Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KM  Glycoprotein; Signal.
FT  SIGNAL 1 21
FT  CHAIN 22 499 ALPHA-AMYLASE.
FT  ACT_SITE 227 227
FT  ACT_SITE 231 231 BY SIMILARITY.
FT  ACT_SITE 251 251 BY SIMILARITY.
FT  ACT_SITE 318 318 BY SIMILARITY.
FT  DISULFID 51 59 BY SIMILARITY.
FT  DISULFID 171 185 BY SIMILARITY.
FT  DISULFID 261 304 BY SIMILARITY.
FT  DISULFID 461 496 BY SIMILARITY.
FT  CARBOHYD 218 218 N-LINKED (GLCNAC... ) (BY SIMILARITY).
SQ  SEQUENCE 499 AA; 54852 MW; 1FB7AE50DA01C03F CRC64;

Query Match 100.0%; Score 44; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.088;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
DB 281 NAFKSTSGS 289

RESULT 5
ID FLGG_BORBU STANDARD; PRT; 265 AA.
AC 051715;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar basal-body rod protein flgg (Distal rod protein).
GN FLGG OR BB0774.
OS Borrelia burgdorferi (Lyme disease spirochete).

```

```

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Castens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kellavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA Van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterback T., Matthey L., McDonald L., Arlrich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi".
RL Nature 390:580-586(1997).
CC -1- SUBUNIT: THE BASAL BODY CONSTITUTES A MAJOR PORTION OF THE
CC FLAGELLAR ORGANELLE AND CONSISTS OF FOUR RINGS (L,P,S, AND M)
CC MOUNTED ON A CENTRAL ROD. THE ROD CONSISTS OF ABOUT 26 SUBUNITS OF
CC FLGG IN THE DISTAL PORTION, AND FLGB, FLGC AND FLGF ARE THOUGHT TO
CC BUILD UP THE PROXIMAL PORTION OF THE ROD WITH ABOUT 6 SUBUNITS
CC EACH (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE FLAGELLA BASAL BODY ROD PROTEINS
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001177; AAC67132.1; -.
DR TIGR: BB0774; -.
DR InterPro: IPR001444; Flag_bb_rod.
DR Pfam: PF00460; flg_bb_rod; 1.
DR PROSITE: PS00588; FLAGELLA_BB_ROD; 1.
KM Flagella; Complete proteome.
SQ SEQUENCE 265 AA; 28960 MW; EEA5E79B8FCBDBD CRC64;

Query Match 72.7%; Score 32; DB 1; Length 265;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
DB 192 NLFKSTSGS 200

RESULT 6
ID SIAP_AERSA STANDARD; PRT; 502 AA.
AC P35823;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE S-layer protein precursor (Tetragonal surface virulence array
DE protein) (A-protein).
GN VAPA.
OS Aeromonas salmonicida.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC Aeromonas.
OX NCBI_TaxID=645;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN=A450;
RX MEDLINE=91332048; PubMed=1869553;
RA Chu S., Cavaignac S., Feutrier J., Phipps B.M., Kostorzynska M.,
RA Kay W.W., Trust T.J.;
RT "Structure of the tetragonal surface virulence array protein and gene

```



```

RT of Aeromonas salmonicida."
RL J. Biol. Chem. 266:15258-15265(1991).
RN [2]
RP SEQUENCE OF 22-48.
RC STRAIN-A400, A450, AND A461:
RX MEDLINE=85124084; PubMed=6525563;
RA Kay W.W., Phipps B.M., Ishiguro E.E., Olafson R.W., Trust T.J.;
RT "Surface layer virulence A-proteins from Aeromonas salmonicida
RL strains."
RL Can. J. Biochem. Cell Biol. 62:1064-1071(1984).
RN [3]
RP SEQUENCE OF 22-48.
RC STRAIN-A450;
RX MEDLINE=83257214; PubMed=6671174;
RA Phipps B.M., Trust T.J., Ishiguro E.E., Kay W.W.;
RT "Purification and characterization of the cell surface virulent A
RT protein from Aeromonas salmonicida."
RL Biochemistry 22:2934-2939(1983).
RN [4]
RP BINDING PROPERTIES.
RX MEDLINE=92112689; PubMed=1730607;
RA Dohg P., Emoedy L., Trust T.J.;
RT "Binding of laminin and fibronectin by the trypsin-resistant major
RT structural domain of the crystalline virulence surface array protein
RT of Aeromonas salmonicida."
RL J. Biol. Chem. 267:43-49(1992).
CC -1- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA. THIS S-LAYER
CC CONFERS VIRULENCE ON THE BACTERIUM INFECTING SALMONID AND NON
CC SALMONID FISH. IT BINDS SPECIFICALLY TO LAMININ AND FIBRONECTIN.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC S-LAYER WITH TETRAGONAL SYMMETRY. TETHERED TO THE OUTER MEMBRANE
CC BY INTERACTION WITH O-SIDE CHAIN OF THE LPS.
CC -1- DOMAIN: THIS S-LAYER PROTEIN IS COMPOSED OF TWO DISTINCT
CC STRUCTURAL DOMAINS: A LARGER MASS TRYPSIN-RESISTANT N-TERMINAL,
CC CONTAINING THE BINDING REGION FOR EXTRACELLULAR MATRIX PROTEINS,
CC AND A SMALLER C-TERMINAL DOMAIN WITH INTERMEDIATE RESISTANCE TO
CC TRYPSIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M64655; AAA21941.1; -.
KW Signal; Cell wall; S-layer; Virulence.
FT SIGNAL 1 21
FT CHAIN 22 502 S-LAYER PROTEIN.
FT VARIANT 26 26 S-> G (IN STRAINS A450 AND A461).
FT VARIANT 35 35 T-> N (IN STRAIN A450).
FT VARIANT 47 47 D-> S (IN STRAINS A400 AND A450).
FT VARIANT 48 48 F-> D (IN STRAIN A400).
SQ SEQUENCE 502 AA; 52868 MW; 8B86AC6C4FE46F77 CRC64;

```

Query Match 72.7%; Score 32; DB 1; Length 502;
 Best Local Similarity 87.5%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 AFASTSGS 9
 ||| ||||
 Db 176 AFASTSGS 183

RESULT 7
 NU98_HUMAN STANDARD: PRT; 920 AA.
 AC P52948;
 DT 01-OCT-1996 (Rel. 34, Created)

```

DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Nuclear pore complex protein Nup98 (Nucleoporin Nup98) (98 kDa
DE nucleoporin).
GN NUP98.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96154188; PubMed=8563754;
RA Borrow J., Shearman A.M., Stanton V.P., Becher R., Collins T.,
RA Williams A.J., Dube I., Katz F., Kwong Y.L., Morris C.,
RA Oiyashiki K., Toyama K., Rowley J., Housman D.E.;
RT "The t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses
RT the genes for nucleoporin NUP98 and class I homeoprotein HOXA9."
RL Nat. Genet. 12:159-167(1996).
CC -1- FUNCTION: PLAYS A ROLE IN THE BIDIRECTIONAL TRANSPORT ACROSS THE
CC NUCLEOPORIN COMPLEX (NPC). THE REPEAT DOMAIN HAS A DIRECT ROLE IN
CC THE TRANSPORT.
CC -1- SUBCELLULAR LOCATION: Nuclear pore complex.
CC -1- DOMAIN: CONTAINS G-L-F-G REPEATS.
CC -1- DISEASE: A FORM OF ACUTE MYELOID LEUKEMIA IS CHARACTERIZED BY A
CC CHROMOSOMAL TRANSLOCATION T(7;11)(P15,P15) WHICH INVOLVES HOXA9
CC AND NUCLEOPORIN NUP98 GENES.
CC -1- SIMILARITY: BELONGS TO THE GLFG FAMILY OF NUCLEOPORINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U41815; AAC50366.1; -.
DR MIM; 601021; -.
DR InterPro: IPR004325; Nucleoporin_FG.
DR Pfam: PF03093; Nucleoporin_FG; 25.
KW Nuclear protein; Transport; Repeat; Chromosomal translocation.
FT DOMAIN 873 877 POLY-GLU.
SQ SEQUENCE 920 AA; 96074 MW; 129DABAFB69253AD CRC64;

```

Query Match 72.7%; Score 32; DB 1; Length 920;
 Best Local Similarity 77.8%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
 ||| ||||
 Db 138 NAFKSTSGS 146

RESULT 8
 ZRF1_MOUSE STANDARD: PRT; 514 AA.
 AC P54103;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Zuo1n related factor-1.
 GN ZRF1 OR DNAC2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96115610; PubMed=8666407;
 RA Hughes R., Chan F.Y., White R.A., Zon L.I.;
 RT "Cloning and chromosomal localization of a mouse cDNA with homology
 RT to the Saccharomyces cerevisiae gene zuo1n.";

```

CC   Genomics 29:546-550(1995).
CC   -1- SUBCELLULAR LOCATION: Nuclear.
CC   -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES.
CC   -1- SIMILARITY: CONTAINS 1 J DOMAIN.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL: U53208; AAC52486.1; -.
DR   HSSP: P25685; 1HDJ.
DR   MGD: MGI:99470; DnaJc2.
DR   InterPro: IPR001623; DnaJ_N.
DR   Pfam: PF00226; DnaJ; 1.
DR   Pfam: PF00249; myb-DNA-binding; 1.
DR   SMART: SM00371; DnaJ; 1.
DR   SMART: SM00395; SART; 1.
DR   PROSITE: PS00636; DnaJ_1; 1.
DR   PROSITE: PS50076; DnaJ_2; 1.
DR   Chapterone: Nuclear Protein.
FT   DOMAIN 94..163 J-DOMAIN.
SQ   SEQUENCE 514 AA; 59510 MW; 10BFD5DFACFB9DE CRC64;
OY   1 NAFKSTGS 9
Db   439 NAFKSTGS 447
Query Match          70.5%; Score 31; DB 1; Length 514;
Best Local Similarity 77.8%; Pred. NO. 50;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

CC   REPEATING STRIPES UNTIL THE END OF THE SHORT PHASE OF GERM-BAND
CC   EXTENSION.
CC   -1- SIMILARITY: TO THE GLI-RELATED GROUP OF C2H2-TYPE ZINC-FINGERS
CC   PROTEINS.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL: X54360; CAA8244.1; -.
DR   PIR: A35817; A35817.
DR   PIR: S12769; S12769.
DR   HSSP: P08151; 2GLI.
DR   Flybase: FBgn0004859; ci.
DR   InterPro: IPR000822; Znf-C2H2.
DR   Pfam: PF00096; Zf-C2H2; 5.
DR   PRINTS: PR00048; ZINC-FINGER.
DR   SMART: SM00355; Znf_C2H2; 5.
DR   PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE: PS50157; ZINC_FINGER_C2H2_2; 4.
DR   Developmental protein; segmentation polarity protein; zinc-finger;
DR   Metal-binding; DNA-binding; Repeat; Nuclear protein.
KW   Metal-binding; DNA-binding; Repeat; Nuclear protein.
FT   DOMAIN 451..603 ZINC_FINGERS.
FT   ZN_FING 451..476 C2H2-TYPE.
FT   ZN_FING 484..511 C2H2-TYPE.
FT   ZN_FING 517..541 C2H2-TYPE.
FT   ZN_FING 547..572 C2H2-TYPE.
FT   ZN_FING 578..603 C2H2-TYPE.
SQ   SEQUENCE 1377 AA; 150881 MW; 85189314D5E7B3B7 CRC64;
OY   1 NAFKSTSG 8
Db   901 NSFSTSG 908
Query Match          70.5%; Score 31; DB 1; Length 1377;
Best Local Similarity 75.0%; Pred. NO. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```
RN [3]
RP SEQUENCE OF 43-227 FROM N.A.
RC STRAIN=129/OLA;
RA Krynetski E.Y., Fessing M.Y., Edick M.J., Evans W.E.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC
CC -1- FUNCTION: CATALYZES THE S-METHYLATION OF THIOPURINE DRUGS SUCH AS
CC 6-MERCAPTOPURINE.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + a thiopurine = S-
CC adenosyl-L-homocysteine + a thiopurine S-methylether.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY. TPMT
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF046887; AAC25919.1; -
DR EMBL; AF037043; AAD02094.1; -
DR EMBL; AF037044; AAD02093.1; -
DR EMBL; AF104832; AAF06075.1; -
DR EMBL; AF104833; AAF06075.1; JOINED.
DR EMBL; AF104826; AAF06075.1; JOINED.
DR EMBL; AF104827; AAF06075.1; JOINED.
DR EMBL; AF104828; AAF06075.1; JOINED.
DR EMBL; AF104829; AAF06075.1; JOINED.
DR EMBL; AF104830; AAF06075.1; JOINED.
DR EMBL; AF104831; AAF06075.1; JOINED.
DR EMBL; AF218593; AAF74424.1; -
DR EMBL; AF218588; AAF74424.1; JOINED.
DR EMBL; AF218589; AAF74424.1; JOINED.
DR EMBL; AF218590; AAF74424.1; JOINED.
DR EMBL; AF218591; AAF74424.1; JOINED.
DR EMBL; AF218592; AAF74424.1; JOINED.
DR MGD; MGI:98812; Tpm.
DR TRANSFERASE: Methyltransferase.
KM VARIANT 69 I -> V (IN STRAIN C57BL/6J).
FT SEQUENCE 240 AA; 27585 MW; 2BA57F30E8EB72D2 CRC64;
SQ
```

Query Match 68.2%; Score 30; DB 1; Length 240;
Best Local Similarity 85.7%; Pred. NO. 37;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 FKSTSGS 9
Db 116 FKSSSGS 122

RESULT 11
TPMT_MUSP
ID TPMT_MUSP STANDARD; PRT; 240 AA.
AC Q9QX22;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Thiopurine S-methyltransferase (EC 2.1.1.67) (Thiopurine
methyltransferase).
DE TPMT.
GN Mus spretus (Western wild mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBT_TaxID=10096;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=EI;
RA Adjei A.A., Johnson G.B., Otterness D.M., Weinsilbom R.M.;
RT "Mouse thiopurine methyltransferase pharmacogenetics: cDNA cloning and

```
RT characterization and processed pseudogene cloning."
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC
CC -1- FUNCTION: CATALYZES THE S-METHYLATION OF THIOPURINE DRUGS SUCH AS
CC 6-MERCAPTOPURINE.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + a thiopurine = S-
CC adenosyl-L-homocysteine + a thiopurine S-methylether.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY. TPMT
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF037045; AAD02094.1; -
DR MGD; MGI:98812; Tpm.
DR TRANSFERASE: Methyltransferase.
KM SEQUENCE 240 AA; 27617 MW; 860A0DC5BD6FBA4C CRC64;
SQ
```

Query Match 68.2%; Score 30; DB 1; Length 240;
Best Local Similarity 85.7%; Pred. NO. 37;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 FKSTSGS 9
Db 116 FKSSSGS 122

RESULT 12
RUVB_RICCN
ID RUVB_RICCN STANDARD; PRT; 342 AA.
AC Q92187;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Holliday junction DNA helicase ruvb.
GN RUVB OR RC0533.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBT_TaxID=781;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii";
RL Science 293:2098-2098(2001).
CC -1- FUNCTION: The ruvb-ruvb complex in the presence of ATP renatures
CC cruciform structure in supercoiled DNA with palindromic sequence,
CC indicating that it may promote strand exchange reactions in
CC homologous recombination. Ruvb is an helicase that mediates the
CC Holliday junction migration by localized denaturation and
CC reannealing (By similarity).
CC -1- SUBUNIT: Forms a complex with ruva (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RUVB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: AE008615; AAL03071.1; ALT_INIT.
 DR InterPro: IPR003959; AAA_ATPase_centre.
 DR InterPro: IPR002182; NB-ARC.
 DR Pfam: PF00004; AAA.1.
 DR Pfam: PF00931; NB-ARC.1.
 DR DNA repair: SOS response; ATP-binding; DNA recombination; Helicase;
 KW complete proteome.
 FT NP_BIND 57 64 ATP (POTENTIAL).
 SQ SEQUENCE 342 AA; 38309 MW; 83891B80C829C28 CRC64;

Query Match 68.2%; Score 30; DB 1; Length 342;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 FRSTSG 8
 |||||
 DB 78 FRSTSG 83

RESULT 13
 E13C_TOBAC STANDARD: PRT; 351 AA.

AC P23432;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Glucan endo-1,3-beta-glucosidase precursor (Ec 3.2.1.39) ((1->3)-beta-glucan endohydrolase) ((1->3)-beta-glucanase) (Beta-1,3-endoglucanase).
 DE SP41A.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=91006039; PubMed=2120041;
 RA Ori N., Sessa G., Lotan T., Himmelhoch S., Fluhr R.;
 RT "A major stylar matrix polypeptide (sp41) is a member of the pathogenesis-related proteins superclass.";
 RL EMBO J. 9:3429-3436(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CV. SAMSUN NN. TISSUE=Style;
 RX MEDLINE=96145508; PubMed=8555460;
 RA Sessa G., Fluhr R.;
 RT "The expression of an abundant transmitting tract-specific endoglucanase (Sp41) is promoter-dependent and not essential for the reproductive physiology of tobacco.";
 RL Plant Mol. Biol. 29:969-982(1995).
 CC -1- FUNCTION: IMPLICATED IN THE DEFENSE OF PLANTS AGAINST PATHOGENS.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,3-beta-D-glucosidic linkages in 1,3-beta-D-glucans.
 CC -1- SUBCELLULAR LOCATION: STYLAR SECRETORY MATRIX.
 CC -1- DEVELOPMENTAL STAGE: MAXIMAL LEVEL OF ACCUMULATION DURING ANTHESIS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 17 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X54430; CAA38302.1; -
 DR EMBL: X81560; CAA57255.1; -
 DR PIR: S12013; S12013.
 DR HSSP: P15737; IGHS.
 DR InterPro: IPR000490; Glyco_hydro_17.

DR Pfam: PF00332; Glyco_hydro_17; 1.
 DR PROSITE: PS00587; GLYCOSYL_HYDROL_F17; 1.
 KW Hydrolyase; Glycosidase; Glycoprotein; signal; Multigene family.
 FT STGNAL 1 32
 FT CHAIN 33 351
 FT MOD_RES 33 33
 FT ACT_SITE 268 268
 FT CARBOHYD 328 328
 FT CARBOHYD 79 79
 FT CARBOHYD 99 99
 FT CARBOHYD 235 235
 FT CARBOHYD 337 337
 SQ SEQUENCE 351 AA; 39186 MW; 588B82C3E8C71F39 CRC64;

Query Match 68.2%; Score 30; DB 1; Length 351;
 Best Local Similarity 66.7%; Pred. No. 55;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MARKSTGS 9
 |||||
 DB 73 NIKFSLNGS 81

RESULT 14
 DPAL_SALTY STANDARD: PRT; 404 AA.

AC P40817;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Diaminopropionate ammonia-lyase (EC 4.3.1.15) (Diaminopropionatase)
 DE (Alpha,beta-diaminopropionate ammonia-lyase).
 GN DPAL OR STM1002.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland W., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;
 RA "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2.";
 RL Nature 413:852-856(2001).
 RN [2]
 RP SEQUENCE OF 1-22: 67-88 AND 239-242.
 RX STRAIN=IFO 12529;
 RX MEDLINE=88087224; PubMed=3275662;
 RA Nagasawa T., Tanizawa K., Satoda T., Yamada H.;
 RT "Diaminopropionate ammonia-lyase from Salmonella typhimurium. Purification and characterization of the crystalline enzyme, and sequence determination of the pyridoxal 5'-phosphate binding peptide.";
 RL J. Biol. Chem. 263:958-964(1988).
 CC -1- FUNCTION: CATALYZES THE ALPHA,BETA-ELIMINATION REACTION OF BOTH L- AND D-ALPHA,BETA-DIAMINOPROPIONATE, THE MOST SUITABLE SUBSTRATES TO FORM PYRUVATE AND AMMONIA. THE L- AND D-ISOMERS OF SERINE ARE ALSO DEGRADED, THOUGH SLOWLY; IT IS THE ONLY SERINE DEHYDRATASE WHICH CAN ELIMINATE AN AMINO GROUP AT THE BETA-CARBON POSITION.
 CC -1- CATALYTIC ACTIVITY: 2,3-diaminopropionate + pyruvate + 2 NH(3).
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- ENZYME REGULATION: COMPETITIVELY INHIBITED BY L- AND D-ALANINE.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: TO E. COLI YGEX AND TO THREONINE DEHYDRATASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE008743; AAL19936.1; -
 DR PIR; A28524; A28524.
 DR StyGene; SG10535; dpaL.
 KW Lyase: Pyridoxal phosphate: Complete proteome.
 FT BINDING 78 78 PYRIDOXAL PHOSPHATE (POTENTIAL).
 FT CONFLICT 16 16 Y -> K (IN REF. 2).
 FT CONFLICT 68 68 E -> G (IN REF. 2).
 FT CONFLICT 242 242 F -> A (IN REF. 2).
 SQ SEQUENCE 404 AA; 44152 MW; 9D605DFEAA691F00 CRC64;

Query Match 68.2%; Score 30; DB 1; Length 404;
 Best Local Similarity 66.7%; Pred. No. 63;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NAKSTSGS 9
 |||||
 Db 75 NAKGIGGS 83

RESULT 15
 FABB_ARATH
 ID FABB_ARATH STANDARD; PRT; 473 AA.
 AC P52410;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 3-oxoacyl-[acyl-carrier-protein] synthase I, chloroplast precursor
 (EC 2.3.1.41) (Beta-ketoacyl-ACP synthase I) (KAS I).
 GN KAS1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID:3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Millar A.A., Kunst L.;
 RT "Isolation of an Arabidopsis cDNA encoding 3-ketoacyl-acyl carrier
 RT protein synthase I.1.";
 RL (In) Plant Gene Register PGR95-027.
 CC -1- FUNCTION: CATALYZES THE CONDENSATION REACTION OF FATTY ACID
 CC SYNTHESIS BY THE ADDITION TO AN ACYL ACCEPTOR OF TWO CARBONS
 CC FROM MALONYL-ACP. SPECIFIC FOR ELONGATION FROM C-10 TO UNSATURATED
 CC C-16 AND C-18 FATTY ACIDS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl]-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + Co(2) +
 CC [acyl-carrier protein].
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: BELONGS TO THE BETA-KETOACYL-ACP SYNTHASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U24177; AAC49118.1; -
 DR HSSP; P39435; IKAS.
 DR InterPro: IPR000794; Ketoacyl-synt.
 DR Pfam: PF00109; ketoacyl-synt; 1.
 DR Pfam: PF02801; ketoacyl-synt_C; 1.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 1.

KW Fatty acid biosynthesis; Transferase; Chloroplast; Transit peptide;
 KW Acyltransferase.
 FT TRANSIT 1 48 CHLOROPLAST (POTENTIAL).
 FT CHAIN 49 473 3-OXOACYL-[ACYL-CARRIER-PROTEIN] SYNTHASE
 FT ACT_SITE 224 224 I.
 SQ SEQUENCE 473 AA; 50351 MW; 51BB5F91A1762008 CRC64;

Query Match 68.2%; Score 30; DB 1; Length 473;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 FKSTSG 8
 |||||
 Db 383 FKSTSG 388

Search completed: June 13, 2002, 09:11:54
 Job time: 1034 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:55:53 ; Search time 73.31 Seconds
(without alignments)
11.797 Million cell updates/sec

Title: US-09-710-339-2_COPY_280_288

Perfect score: 44

Sequence: 1 NAFKSTSGS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	478	2 JK0201	alpha-amyase (EC
2	44	100.0	498	2 A48305	alpha-amyase (EC
3	44	100.0	499	1 ALAS1	alpha-amyase (EC
4	44	100.0	499	1 ALAS1	alpha-amyase (EC
5	44	100.0	499	2 B48305	alpha-amyase (EC
6	44	100.0	499	2 JS0663	alpha-amyase (EC
7	44	100.0	499	2 JT0466	alpha-amyase (EC
8	44	100.0	499	2 JN0588	alpha-amyase (EC
9	35	79.5	469	2 S17726	H-transferring AT
10	34	77.3	102	2 S69741	hypothetical prote
11	33	75.0	297	2 H83888	transmembrane lipo
12	33	75.0	631	2 S72270	alpha-amyase (EC
13	32	72.7	100	2 A72572	hypothetical prote
14	32	72.7	265	2 E70196	flagellar basal bo
15	32	72.7	502	2 A39351	tetragonal surface
16	32	72.7	513	1 S60303	serine/threonine-s
17	32	72.7	681	2 T16353	hypothetical prote
18	32	72.7	1649	2 C86822	hypothetical prote
19	32	70.5	234	2 AG1512	hypothetical prote
20	31	70.5	251	2 AH2520	hypothetical prote
21	31	70.5	419	2 S41607	atrolysin A (EC 3.
22	31	70.5	471	2 B96531	hypothetical prote
23	31	70.5	492	2 F86868	chitinase (EC 3.2.
24	31	70.5	492	2 T47146	hypothetical prote
25	31	70.5	621	2 A57591	Id-associated prot
26	31	70.5	737	2 T02178	hypothetical prote
27	31	70.5	760	2 T41644	hypothetical trp-a
28	31	70.5	766	2 G71437	probable resistanc
29	31	70.5	1041	2 C85189	disease resistance

30	31	70.5	1103	2 A85189	disease resistance
31	31	70.5	1125	2 H87644	Tom-dependent rec
32	31	70.5	1130	2 T21134	hypothetical prote
33	31	70.5	1184	2 H71436	hypothetical prote
34	31	70.5	1256	2 C71436	probable resistance
35	31	70.5	1301	2 D85188	disease resistance
36	31	70.5	1304	2 G85188	disease resistance
37	31	70.5	1317	2 B85189	disease resistance
38	31	70.5	1377	2 A38926	disease resistance
39	31	70.5	1422	2 B71437	DNA-binding protei
40	31	70.5	1484	2 C97196	probable membrane
41	31	70.5	1588	2 A86036	probable resistance
42	31	70.5	1588	2 H91188	probable adhesin Z
43	31	70.5	2088	2 E71436	probable adhesin E
44	31	70.5	2467	2 D71437	hypothetical prote
45	30	68.2	50	2 H70224	probable resistance

ALIGNMENTS

RESULT 1

JK0201

alpha-amyase (EC 3.2.1.1) - Aspergillus oryzae

N:Alternate names: glycoenzyme; Taka-amyase A

C:Species: Aspergillus oryzae

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998

C:Accession: JK0201

R: Toda, H.; Kondo, K.; Narita, K.

Proc. Jpn. Acad. 58B, 208-212, 1982

A:Title: The complete amino acid sequence of Taka-amyase A.

A:Reference number: JK0201

A:Accession: JK0201

A:Molecule type: protein

A:Residues: 1478 <TOD>

C:Comment: One atom of calcium per molecule is essential for the activity.

C:Comment: This enzyme is a glycoprotein.

C:Comment: See also PIR:JT0466 and PIR:JS0240.

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: Aspergillus alpha-amyase; alpha-amyase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:173-300/Domain: alpha-amyase core homology <AMY>

F:197/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:210/230/297/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 44; DB 2; Length 478;

Best Local Similarity 100.0%; Pred. No. 0.19; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9

DB 260 NAFKSTSGS 268

RESULT 2

A48305

alpha-amyase (EC 3.2.1.1) A precursor - Aspergillus awamori

C:Species: Aspergillus awamori

C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997

C:Accession: A48305

R: Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J

Curr. Genet. 17, 203-212, 1990

A:Title: Cloning, characterization, and expression of two alpha-amyase genes from As

A:Reference number: A48305; M01D:90254827

A:Accession: A48305

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1498 <KOR>

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 44; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NAFKSTSGS 9
|||||
Db 281 NAFKSTSGS 289

RESULT 3

ALAS1
alpha-amylase (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae
N:Alternate names: alpha-amylase isozyme II; glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C>Date: 30-Sep-1991 #sequence.revision 30-Sep-1991 #text_change 15-Sep-2000
C:Accession: S04548; A33214; J50240; A91930; A93767; A10627
R:Witsel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
Mol. Microbiol. 3, 3-14, 1989
A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exon
A:Reference number: S04548; MUID:89237897
A:Accession: S04548
A:Molecule type: DNA
A:Residues: 1-499 <MIR>
A:Cross-references: EMBL:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921
A:Genetics: AMY1
A:Accession: A33214
A:Molecule type: mRNA
A:Residues: 1-499 <M12>
A:Cross-references: GB:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921
R:Genes, M.J.; Dove, M.J.; Seligy, V.L.
Gene 79, 107-117, 1989
A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each containing
A:Reference number: J50240; MUID:89378767
A:Accession: J50240
A:Molecule type: DNA
A:Residues: 1-499 <GEN>
A:Genetics: AMY2
A:Note: the authors refer to this as isozyme II
R:Isemura, S.; Ikenaka, T.
J. Biochem. 74, 1-10, 1973
A:Reference number: A91930; MUID:74001521
A:Accession: A91930
A:Molecule type: protein
A:Residues: 206-225 <ISR>
R:Marita, K.
Proc. Jpn. Acad. 51, 285-290, 1975
A:Reference number: A93767
A:Accession: A93767
A:Molecule type: protein
A:Residues: 434-443, 446-447, 'Q', 449-458, 'GTTV', 459-464, 467-468, 'B', 470, 'B', 472-499 <NAR>
R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
J. Biochem. 95, 697-702, 1984
A:Title: Structure and possible catalytic residues of Taka-amylase A.
A:Reference number: A37454; MUID:84212370
A:Contents: annotation; X-ray crystallography, 3.0 angstroms
R:Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkemburg, J.P.; Wilkinson, A.
Submitted to the Brookhaven Protein Data Bank, August 1992
A:Reference number: A51548; PDB:6TAA
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 22-497
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics: <AMY1>
A:Gene: amy1
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Genetics: <AMY1>
A:Gene: amy1
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; po
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase 1 #status experimental <MAT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:51-59,171-185,261-304,461-496/disulfide bonds: #status experimental
F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:227,251,318/Active site: Asp, Glu, Asp #status predicted

Query Match 100.0%; Score 44; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NAFKSTSGS 9
|||||
Db 281 NAFKSTSGS 289

RESULT 4

ALAS3
alpha-amylase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae
N:Alternate names: alpha-amylase isozyme I; glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C>Date: 30-Sep-1991 #sequence.revision 30-Sep-1991 #text_change 18-Jun-1999
C:Accession: S04549; A33215; A44713
R:Witsel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
Mol. Microbiol. 3, 3-14, 1989
A:Title: Three alpha-amylase genes of Aspergillus oryzae exhibit identical intron-exo
A:Reference number: S04548; MUID:89237897
A:Accession: S04549
A:Molecule type: DNA
A:Residues: 1-499 <MIR>
A:Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
A:Accession: A33215
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-499 <M12>
A:Cross-references: GB:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
R:Genes, M.J.; Dove, M.J.; Seligy, V.L.
Gene 79, 107-117, 1989
A:Title: Aspergillus oryzae has two nearly identical Taka-amylase genes, each contain
A:Reference number: J50240; MUID:89378767
A:Accession: A44713
A:Molecule type: DNA
A:Residues: 1-499 <GEN>
A:Note: the authors refer to this as isozyme I
R:Matsuura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
J. Biochem. 95, 697-702, 1984
A:Title: Structure and possible catalytic residues of Taka-amylase A.
A:Reference number: A37454; MUID:84212370
A:Contents: annotation; X-ray crystallography, 3.0 angstroms
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics: AMY1
A:Gene: amy1
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: calcium; carbohydrate digestion; glycoprotein; glycosidase; hydrolase; me
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase 3 #status experimental <MAT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:51-59,171-185,261-304,461-496/disulfide bonds: #status experimental
F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:231,251,318/Active site: His, Glu, Asp #status experimental

Query Match 100.0%; Score 44; DB 1; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
|||||
Db 281 NAFKSTSGS 289

RESULT 5

alpha-amylase (EC 3.2.1.1) B precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: B48305
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Asper
A:Reference number: A48305; MOID:90234827
A:Accession: B48305
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499 <KOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMY>
F:194-321/Domain: alpha-amylase core homology <AMY>

Query Match 100.0%; Score 44; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
|||||
Db 281 NAFKSTSGS 289

RESULT 6

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus sp.
C:Species: Aspergillus sp.
C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
C:Accession: J50663
R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
Biosci. Biotechnol. Biochem. 56, 174-179, 1992
A:Title: Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its expression
A:Reference number: J50663; MOID:92323146
A:Accession: J50663
A:Molecule type: mRNA
A:Residues: 1-499 <SHI>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <ALP>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 44; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
|||||
Db 281 NAFKSTSGS 289

RESULT 7

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: J70466
R:Rada, S.; Iimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
Agric. Biol. Chem. 53, 593-599, 1989
A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Asperg
A:Reference number: J70466
A:Accession: J70466
A:Molecule type: DNA
A:Residues: 1-499 <TAD>
C:Comment: See also PIR:J70201 and PIR:J50240.
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics:
A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <MAT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 100.0%; Score 44; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9
|||||
Db 281 NAFKSTSGS 289

RESULT 8

alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: Taka-amylase A
C:Species: Aspergillus oryzae
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
C:Accession: JN0588
R:Tsuakoshi, N.; Furukawa, M.; Nagaba, H.; Kirita, N.; Tsuboi, A.; Ueda, S.
Gene 84, 319-327, 1989
A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A: evidence for
A:Reference number: JN0588; MOID:90128276
A:Accession: JN0588
A:Molecule type: mRNA
A:Residues: 1-499 <TSU>
C:Comment: The alpha amylases are encoded by multigene family.
C:Genetics:
A:Gene: Taa-G1
A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-499/Product: alpha-amylase #status predicted <MAT>
F:194-321/Domain: alpha-amylase core homology <AMY>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 44; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NAFKSTSGS 9

Db 281 NAFKSTSGS 289

RESULT 9

S17726 H+-transporting ATP synthase (EC 3.6.1.34) beta chain - *Bacillus firmus* (strain OF4)

C:Species: *Bacillus firmus*

C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 19-Jan-2001

C:Accession: S17726

R:Ivey, D.M.; Krulwich, T.A.

Mol. Gen. Genet. 229: 292-300, 1991

A:Title: Organization and nucleotide sequence of the atp genes encoding the ATP synthase

A:Reference number: S17719; M01D:92017655

A:Accession: S17726

A:Molecule type: DNA

A:Residues: 1-469 <1VE>

A:Cross-references: EMBL:M60117; NID:g2988387; PIDN:MAC08044.1; PID:g142551

C:Genetics:

A:Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase al

C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; nucleotide binding

F:157-164/Region: nucleotide-binding motif A (P-loop)

F:183-352/Domain: H+-transporting ATP synthase alpha chain homology <ATP>

Query Match

Best Local Similarity 79.5%; Score 35; DB 2; Length 469;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9

Db 28 NAFKSTSGA 36

RESULT 10

S69741 hypothetical protein YDR455c - yeast (*Saccharomyces cerevisiae*)

C:Species: *Saccharomyces cerevisiae*

C:Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001

C:Accession: S69741

R:Dieterich, F.S.

submitted to the EMBL Data Library, August 1995

A:Description: The sequence of S. cerevisiae lambda 3641 and cosmids 9461, 9831, and 941

A:Reference number: S69555

A:Accession: S69741

A:Molecule type: DNA

A:Residues: 1-102 <DIE>

A:Cross-references: EMBL:U33007; NID:g927685; PID:g2194171; GSPDB:GN00004; MIPS:YDR455c

C:Genetics:

A:Gene: MIPS:YDR455c

A:Map position: 4R

C:Superfamily: *Saccharomyces cerevisiae* hypothetical protein YDR455c

Query Match

Best Local Similarity 77.3%; Score 34; DB 2; Length 102;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 FKSTSGS 9

Db 15 FKSTSGS 21

RESULT 11

H83888 transmembrane lipoprotein lpIC [imported] - *Bacillus halodurans* (strain C-125)

C:Species: *Bacillus halodurans*

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 14-Dec-2001

C:Accession: H83888

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28: 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A:Reference number: A83650; M01D:20512582; PMID:11058132

A:Accession: H83888

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-297 <STO>

A:Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA05631.1; GSPDB:C

A:Experimental source: strain C-125

C:Genetics:

A:Gene: lpIC

C:Superfamily: maltose transport protein malG

Query Match

Best Local Similarity 75.0%; Score 33; DB 2; Length 297;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NAFKSTSG 8

Db 244 NAFKSTSG 251

RESULT 12

S72270 alpha-amylase (EC 3.2.1.1) precursor - *Cryptococcus* sp. (strain CS2)

C:Species: *Cryptococcus* sp.

A:Variety: strain CS2

C:Date: 23-Apr-1998 #sequence_revision 08-May-1998 #text_change 20-Jun-2000

C:Accession: S72270

R:Refuji, H.; Chino, M.; Kato, M.; Iimura, Y.

Biochem. J. 318, 989-996, 1996

A:Title: Raw-starch-digesting and thermostable alpha-amylase from the yeast *Cryptococ*

A:Reference number: S72270; M01D:96433120

A:Accession: S72270

A:Molecule type: DNA

A:Residues: 1-631 <IEF>

A:Cross-references: EMBL:D83540; NID:g1595852; PIDN:BAA12010.1; PID:g1595853

A:Experimental source: strain S-2

C:Genetics:

A:Gene: amy-CS2

A:Introns: 289/2; 326/2

C:Function:

A:Pathway: glycogen/starch degradation

C:Superfamily: *Aspergillus* alpha-amylase; alpha-amylase core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

F:1-20/Domain: signal sequence #status predicted <STG>

F:21-631/Product: alpha-amylase #status predicted <MAT>

F:206-335/Domain: alpha-amylase core homology <AMT>

Query Match

Best Local Similarity 75.0%; Score 33; DB 2; Length 631;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9

Db 295 NAFQSTSGS 303

RESULT 13

A72572 hypothetical protein APE1858 - *Aeropyrum pernix* (strain K1)

C:Species: *Aeropyrum pernix*

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: A72572

R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

DNA Res. 6: 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aero*

A:Reference number: A72450; M01D:99310339

A:Accession: A72572

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-100 <KAW>

A:Cross-references: DDbj:AP000065; MIM:g5105244; PDB:5AA8086.1; PDB:g5105549
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1858
A:Superfamily: Aeropyrum pernix hypothetical protein APE1858

Db 176 AFKLTSGS 183

Search completed: June 13, 2002, 08:55:55
Job time: 375 sec

Query Match	72.7%	Score 32;	DB 2;	Length 100;
Best Local Similarity	77.8%	Pred. No. 14;		
Matches	7;	Conservative	1;	Mismatches
				Indels
				Gaps
QY	1	NAFRTSGS	9	
Db	43	NAFCSSGS	51	

RESULT 14
E70196

flagellar basal body rod protein fligc - Lyme disease
 C/Species: Borrelia burgdorferi (Lyme disease)
 C/Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 20-Aug-1999
 C/Accession: E70196
 R./Rater: C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Whitehagen, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Bowman, C.; Giallard, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.; Nature 390, 580-586, 1997
 A./Authors: Smith, H.O.; Venter, J.C.
 A./Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
 A./Reference number: A70100; MUID:98065943
 A./Accession: E70196
 A./Status: preliminary; nucleic acid sequence not shown; translation not shown
 A./Molecule type: DNA
 A./Residues: 1-265 <1FE>
 A./Cross-references: GB:AE001177; GB:AE000783; MID:92668711; PIDN:AA067132.1; PID:926687Z
 A./Experimental source: strain B31
 C./Genetics:
 A./Gene: fligc
 A./Superfamily: rod protein fligf

Query Match

Query Match	72.7%;	Score 32;	DB 2;	Length 265;
Best Local Similarity	66.7%;	Pred. No. 38;		
Matches	6; Conservative	1; Mismatches	2; Indels	0; Gaps

1 NAFKSTSGS 9

1 NAFKSTSGS 9
y | | | : | |
192 NLFKRETAGS 200
b

RESULT 15

tetragonal surface virulence array protein precursor - *Aeromonas salmonicida*
C:Species: *Aeromonas salmonicida*
C:Date: 27-Mar-1992 #sequence_revision 10-Apr-1992 #text_change 30-Sep-1993
C:Accession: A39351
R:Chu, S.; Cavaignac, S.; Feutrier, J.; Phipps, B.M.; Kostirzyska, M.; Kay, W.W.; Trust,
J. Biol. Chem. 266, 15258-15265, 1991
A:Title: Structure of the tetragonal surface virulence array protein and gene of *Aeromonas*
A:Reference number: A39351; MUID:91332048
A:Accession: A39351
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-502 <CHD>
A:Cross-references: GB:M64655
A:Note: the authors translated the codon CCG for residue 98 as Phe

Query Match

Query Match	72.7%;	Score 32;	DB 2;	Length 502;
Best Local Similarity	87.5%;	Pred. No. 73;		
Matches	7; Conservative	0; Mismatches	1; Indels	0; Gaps

QY 2 AFKSTSGS 9
||| ||||

Fri Jun 14 10:36:47 2002

us-09-710-339-2_copy_280_288.rpr

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:53:28 ; Search time 158.68 Seconds
(without alignments)
6.300 Million cell updates/sec

Title: US-09-710-339-2_COPY_280_288

Perfect score: 44

Sequence: 1 NAFKSTSGS 9

Scoring table:

BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

_A_Geneseq_032802:*

- 1: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/hold-genseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	478	15	AA46065
2	44	100.0	478	16	AA46065
3	44	100.0	478	16	AA46065
4	44	100.0	478	16	AA46065
5	44	100.0	478	16	AA46065
6	44	100.0	478	16	AA46065
7	44	100.0	478	16	AA46065
8	44	100.0	478	16	AA46065
9	44	100.0	478	16	AA46065
10	44	100.0	478	16	AA46065
11	44	100.0	478	16	AA46065

12	32	72.7	411	22	ABG24358	Novel human diago
13	32	72.7	1200	22	ABB64803	Drosophila melanog
14	32	72.7	1712	22	ABB44566	Human wound healin
15	31	70.5	367	21	AA955522	Human cyclic nucle
16	31	70.5	473	22	AA66752	Human ATP-depend
17	31	70.5	473	22	AA678852	Human ATP-depend
18	31	70.5	474	22	AA67531	Human ATP-depend
19	31	70.5	490	21	AA955521	Human cyclic sequenc
20	31	70.5	500	21	AA955521	Human cyclic sequenc
21	31	70.5	576	22	AA667533	Human acid sequenc
22	31	70.5	621	20	AA67029	Breast cancer asso
23	31	70.5	632	17	AA85300	Arabidopsis pathog
24	31	70.5	684	22	AA62678	Human type II phos
25	31	70.5	684	22	AA62678	Human cyclic nucle
26	31	70.5	766	22	AA609526	Novel human diago
27	31	70.5	934	22	AA609526	Human type II phos
28	31	70.5	1269	17	AA603659	RPP5 downy mildew
29	31	70.5	1391	22	AB58810	Drosophila melanog
30	30	68.2	116	18	AA20165	H. pylori flagella
31	30	68.2	116	18	AA24596	H. pylori flagella
32	30	68.2	226	21	AA44938	Artichoke sulphate
33	30	68.2	234	18	AA20722	H. pylori flagella
34	30	68.2	387	21	AA49863	Arabidopsis thalia
35	30	68.2	394	22	AB65646	Drosophila melanog
36	30	68.2	397	22	AB65647	Drosophila melanog
37	30	68.2	429	22	AA59671	B. napus KAS prote
38	30	68.2	431	22	AA59670	B. napus KAS prote
39	30	68.2	441	13	AA22052	A. thaliana KAS pr
40	30	68.2	441	15	AA54027	Brassica synthase
41	30	68.2	455	20	AA537419	Synthase factor B
42	30	68.2	499	21	AA61597	Chlamydia trachoma
43	30	68.2	499	21	AA61597	Arabidopsis thalia
44	30	68.2	513	21	AA627321	Arabidopsis thalia
45	30	68.2	534	21	AA631580	Arabidopsis thalia

ALIGNMENTS

RESULT	1
AA46065	standard; protein; 478 AA.
XX	
XX	
AC	AA46065;
XX	
DF	18-JUL-1994 (first entry)
XX	
DE	Mutant alpha-amyase.
XX	
KW	Methionine substitution; stability; activity; detergent;
KM	dishwashing agents; liquifaction agents.
XX	
OS	Aspergillus oryzae.
XX	
PN	WO9402597-A.
XX	
PD	03-FEB-1994.
XX	
PF	06-JUL-1993; 93WO-DK00230.
XX	
PR	23-JUL-1992; 92DK-0000946.
XX	
PR	16-DEC-1992; 92DK-0001503.
XX	
PR	15-MAR-1993; 93DK-0000292.
XX	
PA	(NOVO) NOVO-NORDISK AS.
XX	
PI	Bisgard-frantzen H, Svendsen A;
XX	
DR	WPI; 1994-048855/06.
XX	
PT	Mutant alpha-amyase from Bacillus species comprising a
PT	methionine substitution - with improved stability and activity at
PT	low pH, for use in detergents, dishwashing agents and

Db 260 nafkstsgs 268

RESULT 4

AAK8270
ID AAK78270 standard; Protein; 478 AA.
XX
AC AAK78270;
XX
DT 17-JUN-1996 (first entry)
XX
DE Aspergillus oryzae alpha amylase (mature protein).
XX
KW Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;
KW starch; thermostable; methionine; Bacillus licheniformis;
KW Bacillus amyloliquefaciens; Bacillus stearothermophilus;
KW Aspergillus oryzae.
XX
OS Aspergillus oryzae.
XX
PN W09521247-A1.
XX
PD 10-AUG-1995.
XX
PE 05-OCT-1994; 94WO-DK00371.
XX
PR 02-FEB-1994; 94DK-0000141.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Marcher D, Nilsson TE, Pedersen HH, Toft AH;
XX
DR WPI; 1995-283767/37.
XX
PT Use of an oxidation stable alpha-amylase - for simultaneous desizing
PT and bleaching or scouring of fabrics conty. starch or starch derivs.
XX
PS Disclosure; Page 25-26; 37pp; English.
XX
CC Oxidation stable alpha amylases can be used for the simultaneous
CC desizing and bleaching or scouring of a fabric comprising starch or
CC starch derivatives. They exhibit a better heat stability,
CC especially in the presence of oxidizing agents. They are obtained
CC from a parent alpha amylase by replacing one or more methionine
CC residues with any amino acid different from Cys or Met, preferably
CC Leu, Thr, Ala, Gly, Ser, Ile or Asp. The parent alpha amylase is
CC pref. derived from a Bacillus species, although alpha amylases of
CC fungal origin can also be used. This sequence is the wild type
CC (unmodified) alpha amylase of Aspergillus oryzae.
XX
SO Sequence 478 AA;

Query Match 100.0%; Score 44; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
DB 260 nafkstsgs 268

RESULT 5
AAW14500
ID AAW14500 standard; protein; 478 AA.
XX
AC AAW14500;
XX
DT 04-JUN-1997 (first entry)
XX
DE Aspergillus oryzae alpha-amylase (mature protein).
XX
KW alpha-amylase; mature; Bacillus licheniformis; termamy1; fungamy1;

KW Aspergillus oryzae; Bacillus amyloliquefaciens; altered property;
KW calcium dependency; substrate binding; stability; pH optimum;
KW thermostability; cleavage; oligosaccharide substrate; dishwashing;
KW washing; detergent additive; fabric desizing; starch liquefaction;
KW sweetener; ethanol production; variant.
XX
OS Aspergillus oryzae.
XX
FH Key Location/Qualifiers
FT Misc-difference 13..45
FT /label= loop 1 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 7-23 of AAW14499 is deleted or
FT replaced with a fragment corresponding to
FT this fragment; claim 33"

Misc-difference 14..40
FT /label= loop 1 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 8-18
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 35" Misc-difference 28..42
FT /label= loop 1 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 12-19 of AAW14499 is deleted or
FT replaced with a fragment corresponding to
FT this fragment; claim 30"

Misc-difference 32..38
FT /label= loop 1 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 14-15
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 32" Misc-difference 66..84
FT /label= loop 2 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 44-57 of AAW14499 is deleted or
FT replaced with a fragment corresponding to
FT this fragment; claim 18"

Misc-difference 70..78
FT /label= loop 2 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 48-51
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 20" Misc-difference 98..210
FT /label= loop 3 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 117-185 of AAW14499 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 24"

Misc-difference 102..206
FT /label= loop 3 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 121-181
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 26" Misc-difference 121..181
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to this fragment is deleted or
FT replaced with a fragment corresponding to
FT 102-206 of AAW14499; claim 41"

Misc-difference 121..174
FT /note= "preferred region where at least one amino acid

RESULT 7
AAU47262 standard; Protein: 75 AA.
AC AUA47262;
DT 27-FEB-2002 (first entry)
DE Propionibacterium acnes immunogenic protein #8158.
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
OS Propionibacterium acnes.
PN WO200181581-A2.
PD 01-NOV-2001.
PF 20-APR-2001; 2001WO-US12865.
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX (CORI-) CORIXA CORP.
XX SKEIXY YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI: 2001-616774/71.
XX N-PSDB: AAS59538.
DR Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Example 1; SEQ ID NO 8457; 1069pp; English.
PS Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 75 AA:
SQ

Query Match 75.0%; Score 33; DB 22; Length 75;
Best Local Similarity 77.8%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 NAFKSTSGS 9
DE ||| | ||| |
DB 54 nafsdsdsgs 62

RESULT 8
ABB67404 standard; Protein: 688 AA.
AC ABB67404;
DT 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 29004.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
OS Drosophila melanogaster.
PN WO200171042-A2.
PD 27-SEP-2001.
PF 23-MAR-2001; 2001WO-US09231.
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li FWD, Myers EW;
XX WPI: 2001-656860/75.
XX N-PSDB: ABL11507.
DR New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 29004; 21pp + Sequence Listing; English.
PS The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB557737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 688 AA:
SQ

Query Match 75.0%; Score 33; DB 22; Length 688;
Best Local Similarity 77.8%; Pred. No. 1,7e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 NAFKSTSGS 9
DE | | | | | | | | | |
DB 385 nlfkstss 393

RESULT 9
ABB66052 standard; Protein: 690 AA.
AC ABB66052;
DT 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 24948.
XX

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI: 2001-656860/75.
DR N-PSDB; ABL10155.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 24948; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABB57737-ABB72072).
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
XX Sequence 690 AA:
SQ
Query Match 75.0%; Score 33; DB 22; Length 690;
Best Local Similarity 77.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 NAFKSTGSG 9
| | | | | | |
Db 387 nlfkstss 395
RESULT 10
AAB43211
ID AAB43211 standard; Protein; 397 AA.
XX
XX AAB43211;
AC
XX 08-FEB-2001 (first entry)
DT
XX
XX Human ORFX ORF2975 polypeptide sequence SEQ ID NO:5950.
DE
XX
XX Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
KW vulnerrary; antiporiatric; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antineoplastic;
KW antiviral; antibacterial; antifungal; antineumatic; antihypertensive;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX

KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
XX Homo sapiens.
OS
XX
XX WO200058473-A2.
PN
XX 05-OCT-2000.
PD
XX 31-MAR-2000; 2000WO-US08621.
XX
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
PA
XX Shinkets RA, Leach M;
PI
XX WPI: 2000-602362/57.
DR N-PSDB; AAC77420.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 11; Page 5133-5134; 5507pp; English.
PS
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytosolic; hepatotropic; vulnerrary;
CC antiporiatric; antiparkinsonian; neurotropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antineoplastic; antibacterial; antiviral; antifungal; antineumatic;
CC antihypertensive; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antineoplastic disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 397 AA:
SQ
Query Match 72.7%; Score 32; DB 21; Length 397;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 NAFKSTG 8
| | | | | | |
Db 77 nlfkstg 84
RESULT 11
ABG08887
ID ABG08887 standard; Protein; 411 AA.
XX
XX ABG08887;
AC
XX 13-FEB-2002 (first entry)
DT
XX
XX Novel human diagnostic protein #8878.
DE
XX

OS	XX	Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX	XX	food supplement; medical imaging; diagnostic; genetic disorder.
PN	XX	Homo sapiens.
PD	XX	WO200175067-A2.
XX	XX	11-OCT-2001.
PF	XX	30-MAR-2001; 2001WO-US08631.
PR	XX	31-MAR-2000; 2000US-0540217.
PR	XX	23-AUG-2000; 2000US-0649167.
PA	XX	(HYSE-) HYSEQ INC.
PI	XX	Dermanac RT, Liu C, Tang YT;
PI	XX	WPI: 2001-639362/73.
DR	XX	N-PsDB; AAS73074.
PT	XX	New isolated polynucleotide and encoded polypeptides, useful in
PT	XX	diagnostics, forensics, gene mapping, identification of mutations
PT	XX	responsible for genetic disorders or other traits and to assess
PT	XX	biodiversity
PS	XX	Claim 20; SEQ ID No 39246; 103bp; English.
CC	XX	The invention relates to isolated polynucleotide (I) and
CC	XX	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	XX	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	XX	and gene mapping, and in recombinant production of (II). The
CC	XX	polynucleotides are also used in diagnostics as expressed sequence tags
CC	XX	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	XX	to restore normal activity of (II) or to treat disease states involving
CC	XX	(II). (II) is useful for generating antibodies against it, detecting or
CC	XX	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	XX	a food supplement. (II) and its binding partners are useful in medical
CC	XX	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	XX	disorders involving aberrant protein expression or biological activity.
CC	XX	The polypeptide and polynucleotide sequences have applications in
CC	XX	diagnostics, forensics, gene mapping, identification of mutations
CC	XX	responsible for genetic disorders or other traits to assess biodiversity
CC	XX	and to produce other types of data and products dependent on DNA and
CC	XX	amino acid sequences. ABG00010-ABG30377 represent novel human
CC	XX	diagnostic amino acid sequences of the invention.
CC	XX	Note: The sequence data for this patent did not appear in the printed
CC	XX	specification, but was obtained in electronic format directly from WIPO
CC	XX	at ftp.wipo.int/pub/Published_pct_sequences.
XX	XX	
SQ	XX	Sequence 411 AA;
OY	XX	1 NAFKSTSGS 9
DB	XX	381 nifstssgs 389
XX	XX	
RESULT 12	XX	
ABG24358	XX	
ID	XX	ABG24358 standard; Protein; 411 AA.
AC	XX	ABG24358;
XX	XX	
DT	XX	18-FEB-2002 (first entry)
XX	XX	
DE	XX	Novel human diagnostic protein #24349.
XX	XX	
XX	XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW		food supplement; medical imaging; diagnostic; genetic disorder.
XX		Homo sapiens.
OS		WO200175067-A2.
PN		11-OCT-2001.
XX		30-MAR-2001; 2001MO-US08631.
XX		31-MAR-2000; 2000US-0540217.
PR		23-AUG-2000; 2000US-0649167.
XX		(HYSE-) HYSEQ INC.
PA		Drmnac RT, Liu C, Tang YT;
XX		WPI: 2001-639362/73.
DR		N-PSTDB; AASB8545.
XX		New isolated polynucleotide and encoded polypeptides, useful in
PT		diagnostics, forensics, gene mapping, identification of mutations
PT		responsible for genetic disorders or other traits and to assess
PT		biodiversity -
PS		Claim 20; SEQ ID No 54717; 103pp; English.
XX		The invention relates to isolated polynucleotide (I) and
CC		polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC		polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC		and gene mapping, and in recombinant production of (II). The
CC		polynucleotides are also used in diagnostics as expressed sequence tags
CC		for identifying expressed genes. (I) is useful in gene therapy techniques
CC		to restore normal activity of (II) or to treat disease states involving
CC		(II). (II) is useful for generating antibodies against it, detecting or
CC		quantitating a polypeptide in tissue, as molecular weight markers and as
CC		a food supplement. (II) and its binding partners are useful in medical
CC		imaging of sites expressing (II). (I) and (II) are useful for treating
CC		disorders involving aberrant protein expression or biological activity.
CC		The polypeptide and polynucleotide sequences have applications in
CC		diagnostics, forensics, gene mapping, identification of mutations
CC		responsible for genetic disorders or other traits to assess biodiversity
CC		and to produce other types of data and products dependent on DNA and
CC		amino acid sequences. ABG00010-ABG30377 represent novel human
CC		diagnostic amino acid sequences of the invention.
CC		Note: The sequence data for this patent did not appear in the printed
CC		specification, but was obtained in electronic format directly from WIPO
CC		at ftp.wipo.int/pub/published_pct_sequences.
XX		Sequence 411 AA:
SQ		
OY	Query Match	72.7%; Score 32; DB 22; Length 411;
Db	Best Local Similarity	66.7%; Pred. NO. 1.6e+02;
	Matches 6; Conservative	2; Mismatches 1; Indels 0; Gaps 0.
	1 NAFKSTSGS 9	
	:	
	381 nlfsssgs 389	
RESULT 13		
ABB64803		
ID	ABB64803 standard; Protein; 1200 AA.	
AC	ABB64803;	
XX		
DT	26-MAR-2002 (first entry)	
DE	Drosophila melanogaster polypeptide SEQ ID NO 21201.	
XX		
XX	Drosophila; developmental biology; cell signalling; insecticide;	
KW	pharmaceutical.	

XX	Drosophila melanogaster.
OS	
XX	W0200171042-A2.
FN	
XX	27-SEP-2001.
PD	
XX	23-MAR-2001; 2001WO-US09231.
Pf	
XX	23-MAR-2000; 2000US-191637P.
PR	
FR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
Pt	Venter JC, Adams M, Li PWD, Myers EW;
DR	
N-PSDB:	ABL08906.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
XX	
PS	Disclosure; SEQ ID NO 21201: 21pp + Sequence Listing: English.
CC	
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16175) and expressed DNA sequences (ABL16176-ABL30511), the encoded proteins (ABB57737-ABB75072).
CC	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at f.p.wipo.int/pub/published_pcr_sequences.
CC	
XX	Sequence 1200 AA:
SQ	
Query Match	72.7%; Score 32; DB 22; Length 1200;
Best Local Similarity	66.7%; Pred. No. 5e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;	
Oy 1 NAFKSTSGS 9 :: :	
Db 1092 sstestsqs 1100	
RESULT 14	
ABBA44566	
ID ABB44566 standard; Protein; 1712 AA.	
NC ABB44566;	
XX	
DT 25-JAN-2002 (first entry)	
XX	
DE Human wound healing related polypeptide SEQ ID NO 23.	
KM Human; mouse; vulnerrary; dermatological; skin disorder; wound healing; gene therapy.	
XX	
OS Homo sapiens.	
XX	
PN CA2325226-A1.	
XX	
PD 17-MAY-2001.	
XX	
PF 16-NOV-2000; 2000CA-2325226.	
XX	
PR 17-NOV-1999; 99DE-1055349.	
PR 17-DEC-1999; 99US-0172511.	
PR 20-JUN-2000; 2000DE-1030149.	

XX	(SWIT-) SWITCH BIOTECH AG.
PA	Reagenbogen J, Wolf E, Coppelt A, Werner S, Halle J;
XX	WPI: 2001-433142/47.
DR	Use of novel polypeptide or its variant or nucleic acid encoding the
PT	polypeptide for diagnosing and/or preventing and/or treating skin
PT	disorders and/or treatment in wound healing or for identifying active
PT	substances
XX	
XX	Disclosure; Page 158-165; 265pp; English.
PS	
XX	The invention relates to the use of a polypeptide (ABA844544-ABA844601,
CC	ABA844606-ABA844623) or its variant or encoding nucleic acid
CC	(ABA81990-ABA81995, ABA82016-ABA82032) with vulnerary and/or
CC	dermatological activity for the diagnosis, prevention and treatment of
CC	skin disorders and treatment in wound healing or for the identification
CC	of pharmacologically active substances. The nucleic acids are useful in
CC	gene therapy.
CC	Note: The printed sequence listing for this specification was incomplete,
CC	terminating part way through SEQ ID NO 106. The remaining data was
CC	obtained from EPO data for an equivalent patent (EP114862).
XX	
SO	Sequence 1712 AA;
XX	
QY	1 NAFKSTSGS 9
Db	138 npfstsgs 146
XX	
RESULT 15	
AAV95522	
ID	AAV95522 standard; Protein; 367 AA.
XX	
AC	AAV95522;
XX	
DT	10-OCT-2000 (first entry)
XX	
DE	Human cyclic nucleotide phosphodiesterase HSPDE10A2.
XX	
KX	Cyclic nucleotide phosphodiesterase; HSPDE10A2; human; cancer;
KW	immune disorder; diagnosis; therapy.
XX	
OS	Homo sapiens.
XX	
FM	Key
FT	Location/Qualifiers
FT	Peptide 88..99
FT	/note="putative cGMP binding motif"
FT	Peptide 260..269
FT	/note="phosphodiesterase signature motif"
PN	WO200040733-A1.
XX	
PD	13-JUL-2000.
XX	
PF	07-JAN-2000; 2000MO-US00371.
XX	
PR	07-JAN-1999; 99US-0226741.
XX	
PA	(INCY-) INCYTE PHARM INC.
XX	
PI	Phillips SC, Ianfear J, Fawcett L, Bandman O, Harrow I;
XX	WPI: 2000-452539/39.
DR	N-PSDB; AAA49973.
XX	

PT New human cyclic nucleotide phosphodiesterases, referred as HSPDE10A1
PT and HSPDE10A2, useful for treating and preventing cancer and immune
PT disorders -
XX
XX
PS
PS

Claim 1; Page 84; 96pp; English.

CC The present sequence is that of HSPDE10A2, a novel human cyclic
CC nucleotide phosphodiesterase. The amino acid sequence was deduced
CC from an isolated CDNA clone (see AAA49973). HSPDE10A2 is a
CC C-terminal splice variant of HSPDE10A1 (see AY95521), being
CC identical between amino acid residues 1-338, but different in the
CC C-terminal portion from residue 339 to 367. It shares 408
CC identity with HSPDE3A1. The invention provides expression vectors,
CC host cells, antibodies, agonists and antagonists, as well as
CC methods for diagnosing, treating or preventing disorders associated
CC with expression of HSPDE10A1, especially cancer and immune
CC disorders.
XX
XX

SQ Sequence 367 AA;

Query Match 70.5%; Score 31; DB 21; Length 367;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NAFKSTSGS 9
|||::|||
Db 269 nafqaksgs 277

Search completed: June 13, 2002, 08:53:29
Job time: 279 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:49:15 ; Search time 57.95 Seconds
(without alignments) 209,904 Million cell updates/sec

Title: US-09-710-339-2

Perfect score: 2684
Sequence: 1 MWAMWSLFLYGLVAPALPA.....LPRVLYPTKRLAGSKICSSS 498

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCRTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2572	95.8	478	1	US-08-720-899-7 Sequence 7, Appli
2	2572	95.8	478	1	US-08-459-610-7 Sequence 7, Appli
3	2572	95.8	478	2	US-08-343-804-7 Sequence 7, Appli
4	2572	95.8	478	2	US-08-600-908A-10 Sequence 10, Appli
5	2572	95.8	478	3	US-08-683-838A-10 Sequence 10, Appli
6	2572	95.8	478	4	US-09-182-859-7 Sequence 7, Appli
7	2503.5	93.3	478	2	US-08-339-715A-2 Sequence 2, Appli
8	1392	51.9	468	1	US-08-470-702-6 Sequence 6, Appli
9	1392	51.9	468	1	US-08-467-831-6 Sequence 6, Appli
10	1391	51.8	468	1	US-08-204-656B-4 Sequence 4, Appli
11	1391	51.8	468	1	US-08-470-702-7 Sequence 7, Appli
12	1391	51.8	468	1	US-08-467-831-7 Sequence 7, Appli
13	1388	51.7	468	1	US-08-204-656B-6 Sequence 6, Appli
14	1388	51.7	468	1	US-08-470-702-8 Sequence 8, Appli
15	1388	51.7	468	1	US-08-467-831-8 Sequence 8, Appli
16	1387	51.7	468	1	US-08-204-656B-2 Sequence 2, Appli
17	1387	51.7	468	1	US-08-204-656B-8 Sequence 8, Appli
18	1387	51.7	468	1	US-08-470-702-9 Sequence 9, Appli
19	1387	51.7	468	1	US-08-467-831-9 Sequence 9, Appli
20	450	16.8	719	4	US-09-386-607-2 Sequence 2, Appli
21	446	16.6	686	3	US-08-947-965-73 Sequence 73, Appli
22	444	16.5	685	3	US-08-947-965-72 Sequence 72, Appli
23	428.5	16.0	655	2	US-08-469-202-28 Sequence 28, Appli
24	428.5	16.0	655	2	US-08-484-434C-35 Sequence 35, Appli
25	427.5	15.9	655	2	US-08-469-202-27 Sequence 27, Appli
26	427.5	15.9	655	2	US-08-484-434C-34 Sequence 34, Appli
27	426.5	15.9	683	3	US-08-947-965-2 Sequence 2, Appli

28	421.5	15.7	624	3	US-08-947-965-78 Sequence 78, Appli
29	418	15.6	676	3	US-08-947-965-71 Sequence 71, Appli
30	413	15.4	686	3	US-08-947-965-70 Sequence 70, Appli
31	413	15.4	687	3	US-08-947-965-75 Sequence 75, Appli
32	404.5	15.1	833	4	US-09-514-302-3 Sequence 3, Appli
33	404.5	15.1	1938	4	US-09-514-302-2 Sequence 2, Appli
34	398	14.8	685	3	US-08-947-965-74 Sequence 74, Appli
35	396	14.8	675	3	US-08-947-965-76 Sequence 76, Appli
36	385.5	14.4	680	3	US-08-947-965-77 Sequence 77, Appli
37	370	13.8	588	2	US-08-339-715A-5 Sequence 5, Appli
38	368	13.7	588	2	US-08-339-715A-1 Sequence 1, Appli
39	367.5	13.7	725	2	US-08-816-105A-4 Sequence 4, Appli
40	367	13.7	588	2	US-08-339-715A-4 Sequence 10, Appli
41	367	13.7	687	1	US-08-470-702-10 Sequence 10, Appli
42	367	13.7	687	1	US-08-467-831-10 Sequence 10, Appli
43	366	13.6	687	1	US-08-204-656B-10 Sequence 3, Appli
44	364	13.6	588	2	US-08-339-715A-3 Sequence 2, Appli
45	259.5	9.7	686	3	US-09-306-922-2 Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-720-899-7
; Sequence 7, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchardt, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720, 899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343, 804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowmeyer Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054, 214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-878-9655
; TELEFAX: 212-878-0123
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-720-899-7

Query Match 95.8%; Score 2572; DB 1; Length 478;
Best local Similarity 99.8%; Pred. No. 1.2e+220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

OY 21 APPADMRQSOSIFELTDRFARTDGSSTATCNTADQKYCGTMOGIIIDKLDYIOGMFTAI 80
      |||||||
Db 1 ATPADMRQSOSIFELTDRFARTDGSSTATCNTADQKYCGTMOGIIIDKLDYIOGMFTAI 60
OY 81 WITPVTAOLPOTTAAGDAHGYWODIYSLNENGTADDLKALSSALHERGMTLMDVVA 140
      |||||||
Db 61 WITPVTAOLPOTTAAGDAHGYWODIYSLNENGTADDLKALSSALHERGMTLMDVVA 120
OY 141 NHMGYDAGSSVDYSEVFKPFSSODYFHPFCFIONYEDOTQVEDCWLGDNTVSLPDLDTTK 200
      |||||||
Db 121 NHMGYDAGSSVDYSEVFKPFSSODYFHPFCFIONYEDOTQVEDCWLGDNTVSLPDLDTTK 180
OY 201 DVYKNEWMDWGSLSVNSISIDGLRIDYKHVKQDFWPGYKKAAGVYICIGEVLDGDPAYTC 260
      |||||||
Db 181 DVYKNEWMDWGSLSVNSISIDGLRIDYKHVKQDFWPGYKKAAGVYICIGEVLDGDPAYTC 240
OY 261 PYQNVMDGVLPYIYPLLNAFKSTSGSMDLYNMINTVKSDCPDSYLLGTFEVNHDPNR 320
      |||||||
Db 241 PYQNVMDGVLPYIYPLLNAFKSTSGSMDLYNMINTVKSDCPDSYLLGTFEVNHDPNR 300
OY 321 FASYTNDIALAKNVAAFIILNDGIPITAYAGOEQHYAGGNDPANREATWLSGYPTDSELYK 380
      |||||||
Db 301 FASYTNDIALAKNVAAFIILNDGIPITAYAGOEQHYAGGNDPANREATWLSGYPTDSELYK 360
OY 381 LTASANAIRNVAISKDTGFVTKKMPYIKDDTTIAMRGTDGSOIYVITLSNKGASGDSYT 440
      |||||||
Db 361 LTASANAIRNVAISKDTGFVTKKMPYIKDDTTIAMRGTDGSOIYVITLSNKGASGDSYT 420
OY 441 LLSLGAGYTAGOQLTEVIGCTTIVTGVSDGNVPVMAGGLPRVLXPTFKLAGSKICSS 498
      |||||||
Db 421 LLSLGAGYTAGOQLTEVIGCTTIVTGVSDGNVPVMAGGLPRVLXPTFKLAGSKICSS 478

RESULT 2
US-08-459-610-7
; Sequence 7, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5801043o No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,610
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054, 214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-459-610-7

Query Match          95.8%; Score 2572; DB 1; Length 478;
Best Local Similarity 99.8%; Pred. No. 1.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 21 APPADMRQSOSIFELTDRFARTDGSSTATCNTADQKYCGTMOGIIIDKLDYIOGMFTAI 80
      |||||||
Db 1 ATPADMRQSOSIFELTDRFARTDGSSTATCNTADQKYCGTMOGIIIDKLDYIOGMFTAI 60
OY 81 WITPVTAOLPOTTAAGDAHGYWODIYSLNENGTADDLKALSSALHERGMTLMDVVA 140
      |||||||
Db 61 WITPVTAOLPOTTAAGDAHGYWODIYSLNENGTADDLKALSSALHERGMTLMDVVA 120
OY 141 NHMGYDAGSSVDYSEVFKPFSSODYFHPFCFIONYEDOTQVEDCWLGDNTVSLPDLDTTK 200
      |||||||
Db 121 NHMGYDAGSSVDYSEVFKPFSSODYFHPFCFIONYEDOTQVEDCWLGDNTVSLPDLDTTK 180
OY 201 DVYKNEWMDWGSLSVNSISIDGLRIDYKHVKQDFWPGYKKAAGVYICIGEVLDGDPAYTC 260
      |||||||
Db 181 DVYKNEWMDWGSLSVNSISIDGLRIDYKHVKQDFWPGYKKAAGVYICIGEVLDGDPAYTC 240
OY 261 PYQNVMDGVLPYIYPLLNAFKSTSGSMDLYNMINTVKSDCPDSYLLGTFEVNHDPNR 320
      |||||||
Db 241 PYQNVMDGVLPYIYPLLNAFKSTSGSMDLYNMINTVKSDCPDSYLLGTFEVNHDPNR 300
OY 321 FASYTNDIALAKNVAAFIILNDGIPITAYAGOEQHYAGGNDPANREATWLSGYPTDSELYK 380
      |||||||
Db 301 FASYTNDIALAKNVAAFIILNDGIPITAYAGOEQHYAGGNDPANREATWLSGYPTDSELYK 360
OY 381 LTASANAIRNVAISKDTGFVTKKMPYIKDDTTIAMRGTDGSOIYVITLSNKGASGDSYT 440
      |||||||
Db 361 LTASANAIRNVAISKDTGFVTKKMPYIKDDTTIAMRGTDGSOIYVITLSNKGASGDSYT 420
OY 441 LLSLGAGYTAGOQLTEVIGCTTIVTGVSDGNVPVMAGGLPRVLXPTFKLAGSKICSS 498
      |||||||
Db 421 LLSLGAGYTAGOQLTEVIGCTTIVTGVSDGNVPVMAGGLPRVLXPTFKLAGSKICSS 478

RESULT 3
US-08-343-804-7
; Sequence 7, Application US/08343804
; Patent No. 5830837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5830837o No. 5830837disk of No. 5830837th America, Inc.
; STREET: 405 Lexington Avenue, 64th floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,804
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 435

```



```

; ATTORNEY/AGENT INFORMATION:
; NAME: Lowmy Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-343-804-7

```

```

Query Match          95.8%; Score 2572; DB 2; Length 478;
Best Local Similarity 99.8%; Pred. No. 1.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 21 ATPADMSOSIYFLITDRFARTDSTTATCNTADOKYCGTMOGIIDKLDYIQMGFTAI 80
DB 1 ATPADMSOSIYFLITDRFARTDSTTATCNTADOKYCGTMOGIIDKLDYIQMGFTAI 60
QY 81 WITPTAQLPQTAYGDAYHGYWODIYSLNENYGTADDLKALSSALHERGMTLWVDVYA 140
DB 61 WITPTAQLPQTAYGDAYHGYWODIYSLNENYGTADDLKALSSALHERGMTLWVDVYA 120
QY 141 NHMGDAGSSVDYSVKFPSSODYFHPFCFIONYEEDOTVEOCWLGDMTVSLPDLDTTK 200
DB 121 NHMGDAGSSVDYSVKFPSSODYFHPFCFIONYEEDOTVEOCWLGDMTVSLPDLDTTK 180
QY 201 DVKKNWYDMWGLSVNSYSDGLRIDTVKHYQKDFWPGYKKAAGVYCIQEVLDGDPAYTC 260
DB 181 DVKKNWYDMWGLSVNSYSDGLRIDTVKHYQKDFWPGYKKAAGVYCIQEVLDGDPAYTC 240
QY 261 PYQNMDCVLNPIYPLPLNAFKSTSGMDLYNMTVKSDDCPDSTLGTFEVENDNPR 320
DB 241 PYQNMDCVLNPIYPLPLNAFKSTSGMDLYNMTVKSDDCPDSTLGTFEVENDNPR 300
QY 321 FASYNDIALAKNVAAFIILNDGIPITAYAGEQHYAGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYNDIALAKNVAAFIILNDGIPITAYAGEQHYAGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNVAISKDGFVYTKNMPYIKDDITIAMKGTDSQIYVILSNKASGDSYT 440
DB 361 LIASANAIRNVAISKDGFVYTKNMPYIKDDITIAMKGTDSQIYVILSNKASGDSYT 420
QY 441 LISGAGYTAGOQLFEVIGCTTVTVGSDGNVPVMAGGLPRVLYPTEKLAGSKICSS 498
DB 421 LISGAGYTAGOQLFEVIGCTTVTVGSDGNVPVMAGGLPRVLYPTEKLAGSKICSS 478

```

```

RESULT 4
; US-08-600-908A-10
; Sequence 10, Application US/08600908A
; Patent No. 5989169
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: "-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5989169 No. 5989169disk of No. 5989169th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,908A
; FILING DATE: 13-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4394.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-600-908A-10

```

```

Query Match          95.8%; Score 2572; DB 2; Length 478;
Best Local Similarity 99.8%; Pred. No. 1.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 21 ATPADMSOSIYFLITDRFARTDSTTATCNTADOKYCGTMOGIIDKLDYIQMGFTAI 80
DB 1 ATPADMSOSIYFLITDRFARTDSTTATCNTADOKYCGTMOGIIDKLDYIQMGFTAI 60
QY 81 WITPTAQLPQTAYGDAYHGYWODIYSLNENYGTADDLKALSSALHERGMTLWVDVYA 140
DB 61 WITPTAQLPQTAYGDAYHGYWODIYSLNENYGTADDLKALSSALHERGMTLWVDVYA 120
QY 141 NHMGDAGSSVDYSVKFPSSODYFHPFCFIONYEEDOTVEOCWLGDMTVSLPDLDTTK 200
DB 121 NHMGDAGSSVDYSVKFPSSODYFHPFCFIONYEEDOTVEOCWLGDMTVSLPDLDTTK 180
QY 201 DVKKNWYDMWGLSVNSYSDGLRIDTVKHYQKDFWPGYKKAAGVYCIQEVLDGDPAYTC 260
DB 181 DVKKNWYDMWGLSVNSYSDGLRIDTVKHYQKDFWPGYKKAAGVYCIQEVLDGDPAYTC 240
QY 261 PYQNMDCVLNPIYPLPLNAFKSTSGMDLYNMTVKSDDCPDSTLGTFEVENDNPR 320
DB 241 PYQNMDCVLNPIYPLPLNAFKSTSGMDLYNMTVKSDDCPDSTLGTFEVENDNPR 300
QY 321 FASYNDIALAKNVAAFIILNDGIPITAYAGEQHYAGNDPANREATWLSGYPTDSELYK 380
DB 301 FASYNDIALAKNVAAFIILNDGIPITAYAGEQHYAGNDPANREATWLSGYPTDSELYK 360
QY 381 LIASANAIRNVAISKDGFVYTKNMPYIKDDITIAMKGTDSQIYVILSNKASGDSYT 440
DB 361 LIASANAIRNVAISKDGFVYTKNMPYIKDDITIAMKGTDSQIYVILSNKASGDSYT 420
QY 441 LISGAGYTAGOQLFEVIGCTTVTVGSDGNVPVMAGGLPRVLYPTEKLAGSKICSS 498
DB 421 LISGAGYTAGOQLFEVIGCTTVTVGSDGNVPVMAGGLPRVLYPTEKLAGSKICSS 478

```

```

RESULT 5
; US-08-683-838A-10
; Sequence 10, Application US/08683838A
; Patent No. 6022724
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg rd-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: "-Amylase Mutants
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6022724 No. 6022724disk of No. 6022724th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor

```

```

; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,838A
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/600,908
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET INFORMATION:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-683-838A-10
```

```

Query Match          95.8%; Score 2572; DB 3; Length 478;
Best Local Similarity 99.8%; Pred. No. 1.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADNRSGSIYFLTDRFARTDGTATCNTADOKYCGGTWOGIIDKLDIYOGMGFTAI 80
DB 1 ATPADNRSGSIYFLTDRFARTDGTATCNTADOKYCGGTWOGIIDKLDIYOGMGFTAI 60

QY 81 WTPVTAQLPOTTAAGDAHGYWODIYSLNENGTADDLKALSALHERGMVLMVDVYA 140
DB 61 WTPVTAQLPOTTAAGDAHGYWODIYSLNENGTADDLKALSALHERGMVLMVDVYA 120

QY 141 NHMGYDAGASSVDYVFEKPFSSQDYFHFPCFIQNYEDQTOVEDCWLGDNTVSLPDLDTTK 200
DB 121 NHMGYDAGASSVDYVFEKPFSSQDYFHFPCFIQNYEDQTOVEDCWLGDNTVSLPDLDTTK 180

QY 201 DYYKNEWYMWGSLVSNYSIDGLRIDTYKHVKQDFWPGYNKAAGYVCIGEVLDDGPAYTC 260
DB 181 DYYKNEWYMWGSLVSNYSIDGLRIDTYKHVKQDFWPGYNKAAGYVCIGEVLDDGPAYTC 240

QY 261 PYQNVMDGYLVNPIYYPPLNFAKSTSGSMDLYNNMINIVKSDCPDSTLLGTFVENHDPNR 320
DB 241 PYQNVMDGYLVNPIYYPPLNFAKSTSGSMDLYNNMINIVKSDCPDSTLLGTFVENHDPNR 300

QY 321 FASYTNDIALAKNVAFIILNDGIPITYAGOBQHYAGSNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAFIILNDGIPITYAGOBQHYAGSNDPANREATWLSGYPTDSELYK 360

QY 381 LIASANAIRNVAISKDTGVTYKKNMPIYKDDTTIAMRGKTDSQIVTITLSKKGASGDSYT 440
DB 361 LIASANAIRNVAISKDTGVTYKKNMPIYKDDTTIAMRGKTDSQIVTITLSKKGASGDSYT 420

QY 441 LSLSGAGYTAGOQLTEVIGCTTVVSGDGNVPVPAAGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGOQLTEVIGCTTVVSGDGNVPVPAAGLPRVLYPTEKLAGSKICSSS 478
```

```

RESULT 6
US-09-182-859-7
; Sequence 7, Application US/09182859
```

```

; Patent No. 6143708
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Jorden
; APPLICANT: Bisgaard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796,204-US
; CURRENT APPLICATION NUMBER: US/09/182,859
; EARLIER FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: 0515/96
; EARLIER FILING DATE: 1996-04-30
; EARLIER APPLICATION NUMBER: 0712/96
; EARLIER FILING DATE: 1996-06-28
; EARLIER APPLICATION NUMBER: 0775/96
; EARLIER FILING DATE: 1996-07-11
; EARLIER APPLICATION NUMBER: 1263/96
; EARLIER FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
; US-09-182-859-7
```

```

Query Match          95.8%; Score 2572; DB 4; Length 478;
Best Local Similarity 99.8%; Pred. No. 1.2e-220;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADNRSGSIYFLTDRFARTDGTATCNTADOKYCGGTWOGIIDKLDIYOGMGFTAI 80
DB 1 ATPADNRSGSIYFLTDRFARTDGTATCNTADOKYCGGTWOGIIDKLDIYOGMGFTAI 60

QY 81 WTPVTAQLPOTTAAGDAHGYWODIYSLNENGTADDLKALSALHERGMVLMVDVYA 140
DB 61 WTPVTAQLPOTTAAGDAHGYWODIYSLNENGTADDLKALSALHERGMVLMVDVYA 120

QY 141 NHMGYDAGASSVDYVFEKPFSSQDYFHFPCFIQNYEDQTOVEDCWLGDNTVSLPDLDTTK 200
DB 121 NHMGYDAGASSVDYVFEKPFSSQDYFHFPCFIQNYEDQTOVEDCWLGDNTVSLPDLDTTK 180

QY 201 DYYKNEWYMWGSLVSNYSIDGLRIDTYKHVKQDFWPGYNKAAGYVCIGEVLDDGPAYTC 260
DB 181 DYYKNEWYMWGSLVSNYSIDGLRIDTYKHVKQDFWPGYNKAAGYVCIGEVLDDGPAYTC 240

QY 261 PYQNVMDGYLVNPIYYPPLNFAKSTSGSMDLYNNMINIVKSDCPDSTLLGTFVENHDPNR 320
DB 241 PYQNVMDGYLVNPIYYPPLNFAKSTSGSMDLYNNMINIVKSDCPDSTLLGTFVENHDPNR 300

QY 321 FASYTNDIALAKNVAFIILNDGIPITYAGOBQHYAGSNDPANREATWLSGYPTDSELYK 380
DB 301 FASYTNDIALAKNVAFIILNDGIPITYAGOBQHYAGSNDPANREATWLSGYPTDSELYK 360

QY 381 LIASANAIRNVAISKDTGVTYKKNMPIYKDDTTIAMRGKTDSQIVTITLSKKGASGDSYT 440
DB 361 LIASANAIRNVAISKDTGVTYKKNMPIYKDDTTIAMRGKTDSQIVTITLSKKGASGDSYT 420

QY 441 LSLSGAGYTAGOQLTEVIGCTTVVSGDGNVPVPAAGLPRVLYPTEKLAGSKICSSS 498
DB 421 LSLSGAGYTAGOQLTEVIGCTTVVSGDGNVPVPAAGLPRVLYPTEKLAGSKICSSS 478
```

```

RESULT 7
US-08-339-715A-2
; Sequence 2, Application US/08339715A
; Patent No. 5965442
; GENERAL INFORMATION:
; APPLICANT: Kaneko, Hiroki
; APPLICANT: Takada, Toshikazu
; APPLICANT: Shimada, Jiro
; APPLICANT: Kuriki, Takashi
; APPLICANT: Yanase, Michiyo
```

APPLICANT: Takata, Hiroki
APPLICANT: Okada, Shigetaka
TITLE OF INVENTION: METHOD OF ALTERING ENZYMES AND A NOVEL
TITLE OF INVENTION: NEOPULULANASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: POBHAM, HAIK, SCHNORRICH & KAUFMAN, LTD.
STREET: Metropolitan Square Building, Suite 800, 1450
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/339,715A
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 306096/1993
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Meiser, Allen S.
REGISTRATION NUMBER: 27,215
REFERENCE/DOCKET NUMBER: 18335,009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-824-8000
TELEFAX: 202-824-8199
TELEX: 248516
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-339-715A-2

Query Match 93.3%; Score 2503.5; DB 2; Length 478;
Best Local Similarity 97.9%; Pred. No. 1.5e-214;
Matches 468; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

21 ATPADWRSQSIYFLTLDFARFDGTTATCNTADOKTGGTMOGIIIDKLDYIYQGMGFTAI 80
1 ATPADWRSQSIYFLTLDFARFDGTTATCNTADOKTGGTMOGIIIDKLDYIYQGMGFTAI 60
81 WITPVTAQLPOTTAAGDAGHYGYWODIYSLNENYGTADDLKALSALHERGMWLVADVA 140
61 WITPVTAQLPOTTAAGDAGHYGYWODIYSLNENYGTADDLKALSALHERGMWLVADVA 120
141 NHMGYDAGSSVDYVYFKPFSSQDYFHPFCFIONYEDQTOVEDCWLGDNTVSLPDLDTTK 200
121 NHMGYDAGSSVDYVYFKPFSSQDYFHPFCFIONYEDQTOVEDCWLGDNTVSLPDLDTTK 180
201 DVYKNEWDWGSIVSNYSIDGLRIDTVKHYOKDFWPGYKNAAGYCTGCEVLDDGPATYC 260
181 DVYKNEWDWGSIVSNYSIDGLRIDTVKHYOKDFWPGYKNAAGYCTGCEVLDDGPATYC 240
261 PYQNMVGVLANPIYPIPLNAFKSTSGSMDDLNNINTVKSDCPSDTLGTFFVENHNDPR 320
241 PYQNMVGVLANPIYPIPLNAFKSTSGSMDDLNNINTVKSDCPSDTLGTFFVENHNDPR 300
321 FASYNDIALAKNVAFTIILNDGPIIYAGQEHYAGGNDPANREATWLSGYPTDSELYK 380
301 FASYNDIALAKNVAFTIILNDGPIIYAGQEHYAGGNDPANREATWLSGYPTDSELYK 360
381 LIASANAIRNVAISKDQGFVYTKNMPLYKDDTTIAMRGSTGSOIYVTLSSKKGSGDSYT 440

Db 361 LIASANAIRNVAISKDQGFVYTKN-PIKDDTTIAMRGSTGSOIYVTLSSKKGSGDSYT 419
QY 441 LLSAGACTAQQOLEVYIGCTTVVSGSNVPVMAGGLPRVLYPTEKLAGSKICSSS 498
Db 420 LLSGASYTAQQOLEVYIGCTTVVSGSNVPVMAGGLPRVLYPTEKLAGSKICSSS 477

RESULT 8
US-08-470-702-6
Sequence 6, Application US/08470702
Patent No. 5631149
GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAKI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,702
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-470-702-6

Query Match 51.9%; Score 1392; DB 1; Length 468;
Best Local Similarity 54.4%; Pred. No. 1.2e-115;
Matches 259; Conservative 81; Mismatches 126; Indels 10; Gaps 5;

22 TPAD-WRSQSIYFLTLDFARFDGTTATCNTADOKTGGTMOGIIIDKLDYIYQGMGFTAI 80
2 TPADWRSQSIYFLTLDFARFDGTTATCNTADOKTGGTMOGIIIDKLDYIYQGMGFTAI 61
81 WITPVTAQLPOTTAAGDAGHYGYWODIYSLNENYGTADDLKALSALHERGMWLVADVA 140
62 WISPYVENITDNATYAGYHGFPMKNITYKINENFTADDLKSLAQLERHDMILWADIVT 121
141 NHMGYDAGSSVDYVYFKPFSSQDYFHPFCFIONYEDQTOVEDCWLGDNTVSLPDLDTTK 200

Db	122	NHYSGDSGSDSIDVSELTPEFNDCXFFHNHYCLISYDDQAOVQOSOMEGSDSSALPDLTATED	18
Qy	201	DVYKNEYMDWGLSYNSYSDIGLRIDYKXHQKDFMFGYKKNKAAGYVIGYVLDGDPAYTC	266
Db	182	SDVASVFSNKKDFGVGNSIDIGLHSDSKKHHDGQGFPEPPFVSASGVYVGEFQDDPAYTC	241
Qy	261	PYQVNMDCVLVYPIYPLPLNAFKSTGSGMDLVMKIMITYKSDCPDSTLLGTFVEVHNMPR	320
Db	242	PYQVNIPEGVSNPIYPTPTFPFKTTDSSSELTQMSIVASASCSDPPLLNFVEVHNENR	301
Qy	321	FASYNIDIALKNNAAFTIILNDGIPITAYGQGHYAGSGNPNANEAUWLSGYPDPSLTK	380
Db	302	FASMTSDSLISMAIAFLLDGGIPVITYGQGLSGSDPNNEEALMLSLCYNKESDYTK	361
Qy	381	LIASANARNAISKDTEFVYTKMWPILKDDTITAMRKGTGSGQSLVILSNKSGASDSTY	440
Db	362	LIANKANARNAAVYODSSYATSOISVIFSNDAVIATKRGSD---VVSFNNLIGSSGSS-D	416
Qy	441	LSLSAGTAAQQLTEVIGCTTIVYVSGSNVPVYMAGLGPVLVLPTEKLSGRKTS	496
Db	417	VTISMTGSSGDLVEVLTGCVSGSSDSD---LQVSTGGGQDIPPIEYPAK---YASDICS	466

```

1      RESULT      9
2      US-08-467-831-6
3      ; Sequence 6, Application US/08467831
4      ; Patent No. 5635378
5      ; GENERAL INFORMATION:
6      ; APPLICANT: MATSUI, IKUO
7      ; APPLICANT: ISHIKAWA, KAZUHIKO
8      ; APPLICANT: MIYAIRI, SACHIO
9      ; APPLICANT: HONDA, KOICHI
10     ; TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
11     ; TITLE OF INVENTION: VARIANT GENE OF THE ENZYME FOR PRODUCING
12     ; TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
13     ; NUMBER OF SEQUENCES: 17
14     ; CORRESPONDENCE ADDRESS:
15     ; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
16     ; STREET: 8110 Gatehouse Road, Suite 500 East
17     ; CITY: Falls Church
18     ; STATE: Virginia
19     ; COUNTRY: U.S.A.
20     ; ZIP: 22042
21     ; COMPUTER READABLE FORM:
22     ; MEDIUM TYPE: Floppy disk
23     ; COMPUTER: IBM PC compatible
24     ; OPERATING SYSTEM: PC-DOS/MS-DOS
25     ; SOFTWARE: PatentIn Release #1.0, Version #1.25
26     ; CURRENT APPLICATION DATA:
27     ; APPLICATION NUMBER: US/08/467,831
28     ; FILING DATE: 06-JUN-1995
29     ; CLASSIFICATION: 435
30     ; PRIOR APPLICATION DATA:
31     ; APPLICATION NUMBER: US 08/204,656
32     ; FILING DATE: 02-MAR-1994
33     ; ATTORNEY/AGENT INFORMATION:
34     ; NAME: WEINER, MARC S.
35     ; REGISTRATION NUMBER: 32,181
36     ; REFERENCE/DOCKET NUMBER: 234-252P
37     ; TELECOMMUNICATION INFORMATION:
38     ; TELEPHONE: (703) 205-8000
39     ; TELEFAX: (703) 205-8050
40     ; TELEX: 248345
41     ; INFORMATION FOR SEQ ID NO: 6:
42     ; SEQUENCE CHARACTERISTICS:
43     ; LENGTH: 468 amino acids
44     ; TYPE: amino acid
45     ; STRANDEDNESS: single
46     ; TOPOLOGY: linear
47     ; MOLECULE TYPE: protein
48     ; HYPOTHETICAL: NO
49     ; ANTI-SENSE: NO
50     ;
51     ; US-08-467-831-6

```

Query Match	51.9%;	Score 1392;	DB 1;	Length 468;
Best Local Similarity	54.4%;	Pred. No. 1.2e-115;		
Matches 259;	Conservative 81;	Mismatches 126;	Indels 10;	Gaps 5.

Qy	22	TPAD--WRSOIYFELTDEFAFRTDSTATONVADOKYCCGWTGGLIDLVIOQMGFTAI	80
Db	2	TNADKRRSOSIQIYQTDFAFRTDSDTSASONTDRLYCGGSPGIGIKKLDYIKDMGFTAI	61
Qy	81	WTPVTAQLPQTAVGAIHBITWOODIYSLINENYKIGTADDLKALSALHERGMTLVMDVVA	140
Db	62	WSPVVENIPDNTAVGAYVHYEFMMKNILYKILNENFGTADDLKSLAEDLRDMLLVMDIVT	121
Qy	141	NHNGYCGAGSSVDYVNEAFPSODYFHPFQFIONYEDQYVDECMJGDNVSLPDLPTTK	200
Db	122	NHFGSDGSDSIDIYSELYPEFDOKYFHNHYCLISNYDDQAKOVOSCEBGSDVALDPLRTED	181
Qy	201	DVKNEMMYMWSLSVNSYIDGLRIDLTVKHYOKDFWPGINKAGVYCIQEYLDSDPAYTC	260
Db	182	SDVASVFNBMWMDPYGNTSIDGLRIDSARKHVDGFFDFEVSAGYVSEVFGQDPATYC	241
Qy	261	PYONVMDGVLPNTPIYEPILNAFKSTSGSMDLYNMINTYKSDCPSDSTLLGTFVYEHNNPR	320
Db	242	PYONYIPGVSNPLPYPTTRFFKTTDSSSLBQMLISVASSCSODPTLLTFNVEHNDHER	301
Qy	321	FASYTNDILAKNVAAFIILNDGFIPIYAGQEGHYAGANDPANREATWLSGYPIDSELYK	380
Db	302	FASWTSDDOSIISNALAFVLLDGDIPVILYIQEGDGLSKSDPNNRKELMLSGYNKESDYK	361
Qy	381	LIASANAIHNYAISKDTGFTYKRWPIYKDDTTIAMRKGTDSQIYTLISLNGASGDSYT	440
Db	362	LIAKANAAHNAAYQDSYATISQSLVSFNSDNHVIATKRS---VYSVFNNIGSSGS-D	416
Qy	441	LSLSGAGYTAGOOLIEVIGCTIYVYMGSDGNVPMGAGPRLVYLPERKLAGKIS 496	
Db	417	VTLISNNGYSSGDEVEVLTCSTVSSSDS--LQVSIQGGQQLIVPVRK--YASDICS 468	

RESULT 10
 US-08-204-656B-4
 ; Sequence 4, Application US/08204656B
 ; Patent No. 5538982
 ; GENERAL INFORMATION:
 ; APPLICANT: Matsui, Ikuo
 ; APPLICANT: Ishikawa, Kazuhiko
 ; APPLICANT: Miyairi, Sachio
 ; APPLICANT: Honda, Koichi
 ; TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
 ; TITLE OF INVENTION: Variant Gene Of The Enzyme And Method For Producing
 ; TITLE OF INVENTION: Oligosaccharide Using The Enzyme
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
 ; STREET: 8110 Gatehouse Road, Suite 500 East
 ; CITY: Falls Church
 ; STATE: Virginia
 ; COUNTRY: U.S.A.
 ; ZIP: 22042
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/204,656B
 ; FILING DATE: 02-MAR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Welner, Marc S.
 ; REGISTRATION NUMBER: 32,181
 ; REFERENCE/DOCKET NUMBER: 224-252P
 ; TELECOMMUNICATION INFORMATION:
 ;

TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
TOPOLOGY: 1 linear
MOLECULE TYPE: protein
US-08-204-6568-4

Query Match 51.8%; Score 1391; DB 1; Length 468;
Best Local Similarity 54.4%; Pred. No. 1.5e-115;
Matches 259; Conservative 81; Mismatches 126; Indels 10; Gaps 5;

QY 22 TPAD-WRSQSIYFLTDREPARTDGTATCMTADOKYCGTMOGIIIDKLDYIQMGFTAI 80
DB 2 TNADKRSQSIYQIYTDREPARTDGTASCMTEDELRYCGSGFGIIRKLDYIKDMGFTAI 61
QY 81 WTPPTAQLPQTAYAGDAHGYWODIYSLNENYGTADLKLALSALHERGMTLWADVYA 140
DB 62 WISPVVENIPDNTAIGAYHGMWKNYIKINENFGTADLKLALDELHDMMLWADVIVT 121
QY 141 NHMGYDAGSSVDYSVEKPFSSQDYFHPFCFIQYVEDQYVEDQWLGDNVSLPDLDTK 200
DB 122 NHYGSQSGSDSIDSEYRPFENDOKYFHNKYCLISNVDQAOVQSCWEGDSSVALPDLRTED 181
QY 201 DVKKNEMTWVGSVLSNYSIDGLRIDTVKHVQKDFWPGYNKAGVYCIGEVLGDPAYTC 260
DB 182 SDVASVFNWVKDFVGNYSIDGLRIDSAKHVDQGFPPDFVASGVYSGVEVFGQDPAVTC 241
QY 261 PYONVMDVLANPYIYPLNLFKSTSGSMDLNMINTVKSDCDDSTLLGTFVENHNDPR 320
DB 242 PYONIPGVSNPPLYPTTRFFKTTDSSSELQMISSVASSCSDPTLLTNFVENHNDPR 301
QY 321 FASYTNDIALAKNVAFIILNDGIPITYAGOEQHYAGGNDPANREATWLSGPTDSELYK 380
DB 302 FASMTSDQSLISMAIAFVLGDIPIVYIGQDGLSGKDPNNREALMLSTGNKESDYK 361
QY 381 LIASNAIRNVAISKDGTGVYTKMMPYIKDDTTAMRKGTDSQIVTLLSNKAGSDSYT 440
DB 362 LIAKANAARNAVAYODSSYATSQLSVIFSNHVIATKRGSGSSD-VVSVFNNLSSGSS-D 416
QY 441 LSLSGAGYTAGOOLTEYIGCTTVYVSGDNVPMAGLPRVLYPTEKLAGSKTCS 496
DB 417 VTISNTGSSGDELVEVLTCTVSGSSD--LQVSIQGGQPOLIYPAK--YASDICS 468

RESULT 11
US-08-470-702-7
Sequence 7, Application US/08470702
Patent No. 5631149
GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO
APPLICANT: HONDA, KOICHI
TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,702
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,656
FILING DATE: 02-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 234-252P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
AMTI-SENSE: NO
US-08-470-702-7

Query Match 51.8%; Score 1391; DB 1; Length 468;
Best Local Similarity 54.4%; Pred. No. 1.5e-115;
Matches 259; Conservative 81; Mismatches 126; Indels 10; Gaps 5;

QY 22 TPAD-WRSQSIYFLTDREPARTDGTATCMTADOKYCGTMOGIIIDKLDYIQMGFTAI 80
DB 2 TNADKRSQSIYQIYTDREPARTDGTASCMTEDELRYCGSGFGIIRKLDYIKDMGFTAI 61
QY 81 WTPPTAQLPQTAYAGDAHGYWODIYSLNENYGTADLKLALSALHERGMTLWADVYA 140
DB 62 WISPVVENIPDNTAIGAYHGMWKNYIKINENFGTADLKLALDELHDMMLWADVIVT 121
QY 141 NHMGYDAGSSVDYSVEKPFSSQDYFHPFCFIQYVEDQYVEDQWLGDNVSLPDLDTK 200
DB 122 NHYGSQSGSDSIDSEYRPFENDOKYFHNKYCLISNVDQAOVQSCWEGDSSVALPDLRTED 181
QY 201 DVKKNEMTWVGSVLSNYSIDGLRIDTVKHVQKDFWPGYNKAGVYCIGEVLGDPAYTC 260
DB 182 SDVASVFNWVKDFVGNYSIDGLRIDSAKHVDQGFPPDFVASGVYSGVEVFGQDPAVTC 241
QY 261 PYONVMDVLANPYIYPLNLFKSTSGSMDLNMINTVKSDCDDSTLLGTFVENHNDPR 320
DB 242 PYONIPGVSNPPLYPTTRFFKTTDSSSELQMISSVASSCSDPTLLTNFVENHNDPR 301
QY 321 FASYTNDIALAKNVAFIILNDGIPITYAGOEQHYAGGNDPANREATWLSGPTDSELYK 380
DB 302 FASMTSDQSLISMAIAFVLGDIPIVYIGQDGLSGKDPNNREALMLSTGNKESDYK 361
QY 381 LIASNAIRNVAISKDGTGVYTKMMPYIKDDTTAMRKGTDSQIVTLLSNKAGSDSYT 440
DB 362 LIAKANAARNAVAYODSSYATSQLSVIFSNHVIATKRGSGSSD-VVSVFNNLSSGSS-D 416
QY 441 LSLSGAGYTAGOOLTEYIGCTTVYVSGDNVPMAGLPRVLYPTEKLAGSKTCS 496
DB 417 VTISNTGSSGDELVEVLTCTVSGSSD--LQVSIQGGQPOLIYPAK--YASDICS 468

RESULT 12
US-08-467-831-7
Sequence 7, Application US/08467831
Patent No. 5635378
GENERAL INFORMATION:
APPLICANT: MATSUI, IKUO
APPLICANT: ISHIKAWA, KAZUHIKO
APPLICANT: MIYAIRI, SACHIO

```

? APPLICANT: HONDA, KOICHI
? TITLE OF INVENTION: VARIANT-TYPE CARBOHYDRATE HYDROLASE,
? TITLE OF INVENTION: VARIANT GENE OF THE ENZYME AND METHOD FOR PRODUCING
? TITLE OF INVENTION: OLIGOSACCHARIDE USING THE ENZYME
? NUMBER OF SEQUENCES: 17
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
? STREET: 8110 Gatehouse Road, Suite 500 East
? CITY: Falls Church
? STATE: Virginia
? COUNTRY: U.S.A.
? ZIP: 22042
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/467,831
? FILING DATE: 06-JUN-1995
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/204,656
? FILING DATE: 02-MAR-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: WEINER, MARC S.
? REGISTRATION NUMBER: 32,181
? REFERENCE/DOCKET NUMBER: 234-252P
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703) 205-8000
? TELEFAX: (703) 205-8050
? TELEX: 248345
? INFORMATION FOR SEQ. ID NO: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 468 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHEICAL: NO
? ANTI-SENSE: NO
? US-08-467-831-7

```

```

Query Match 51.8%; Score 1391; DB 1; Length 468;
Best Local Similarity 54.4%; Pred. No. 1.5e-115;
Matches 259; Conservative 81; Mismatches 126; Indels 10; Gaps 5;

QY 22 TPAD-WRSQSIYFLTDRFARTDGTATCNTADOKYCGGTGOGILDKDYIQGMGFTAI 80
DB 2 TNADKMRQSIOYITDFARTDGTASCTEDRIKYGSGFQGIKKDKMGFTAI 61
QY 81 WTPVTAQLPOTAYAGYHGYWOODIYSLNENYGTADDLKALSLAHERGMYLMDVVA 140
DB 62 WISPVENIPNTAYAGYHGMKMKNIYKINENFTADDLKSLAOELHDMMLMDVIVT 121
QY 141 NHMGYDAGSSVDYSEVFEKPFSSQDYFHPFCFIONYEQOTQVEDCWLIGNTVSLPDLITTK 200
DB 122 NHGSDGSGSDIYSEVTEPFNDQKYFHNCLISNDDAQVOQSCWEGSDVALPLDITTED 181
QY 201 DVKKREMYDMVGLSVNSYSDGLRIDTVKHYOKDFWPGYNKAAGYTCIGEVLDGDPAYTC 260
DB 182 SDVASVFNKDYKDFVGNYSIDGLRIDSAKHVDGFFPFVSAAGYVSGEVFGQDPAYTC 241
QY 261 PYQNVMDGVNPIYIYPLNAFKSTGSGMDLYNNMINTVKSDCDPSDSTLLGTFFVENHDPN 320
DB 242 PYQNVIPGSNNPLYPTTRFFKTTDSSSELQWISSVASSCDPILLTFNFEVNHDER 301
QY 321 FASYTNDIALAKNVAFTIINDGIPITYAGQOHYAGNDPANDPANTATLSGYPTDSLEYK 380
DB 302 FASMTSDGLSLNAIAFVLLGDGIPVITYGQGLSGKSDPNRPAALMSGYNKESDYYK 361
QY 381 LIASNAIRNVAISKDTGFVYTKNMPYKDDPTTIAMRGTDGSOQVITLLSKGASGSYTC 440

```

```

DB 362 LIAKNAARNAVAYDSSVATSOALSIFSNHVIATKRGSG---VVSVFNMIGSSGSS-D 416
QY 441 LSLGAGTATGACQQLTEVIGCTTYVSGDGNVPVPAGGLPVLVITPELASKICS 496
DB 417 VTIWNTGYSGEDIVEVLTCSIVSGSSD--LQVSTIGSGQFOIFVPAK--VYSDICS 468

RESULT 13
US-08-204-656B-6
? Sequence 6, Application US/08204656B
? Patent No. 5538882
? GENERAL INFORMATION:
? APPLICANT: Matsui, Ikuro
? APPLICANT: Ishikawa, Kazuhiko
? APPLICANT: Miyairi, Sachio
? APPLICANT: Honda, Koichi
? TITLE OF INVENTION: Variant-Type Carbohydrate Hydrolase,
? TITLE OF INVENTION: Variant Gene of The Enzyme And Method For Producing
? NUMBER OF SEQUENCES: 18
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
? STREET: 8110 Gatehouse Road, Suite 500 East
? CITY: Falls Church
? STATE: Virginia
? COUNTRY: U.S.A.
? ZIP: 22042
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/204,656B
? FILING DATE: 02-MAR-1994
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Weiner, Marc S.
? REGISTRATION NUMBER: 32,181
? REFERENCE/DOCKET NUMBER: 234-252P
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703) 205-8000
? TELEFAX: (703) 205-8050
? TELEX: 248345
? INFORMATION FOR SEQ. ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 468 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-204-656B-6

Query Match 51.7%; Score 1388; DB 1; Length 468;
Best Local Similarity 54.4%; Pred. No. 2.8e-115;
Matches 259; Conservative 80; Mismatches 127; Indels 10; Gaps 5;

QY 22 TPAD-WRSQSIYFLTDRFARTDGTATCNTADOKYCGGTGOGIITDKDYIQGMGFTAI 80
DB 2 TNADKMRQSIOYITDFARTDGTASCTEDRIKYGSGFQGIKKDKMGFTAI 61
QY 81 WTPVTAQLPOTAYAGYHGYWOODIYSLNENYGTADDLKALSLAHERGMYLMDVVA 140
DB 62 WISPVENIPNTAYAGYHGMKMKNIYKINENFTADDLKSLAOELHDMMLMDVIVT 121
QY 141 NHMGYDAGSSVDYSEVFEKPFSSQDYFHPFCFIONYEQOTQVEDCWLIGNTVSLPDLITTK 200
DB 122 NHGSDGSGSDIYSEVTEPFNDQKYFHNCLISNDDAQVOQSCWEGSDVALPLDITTED 181
QY 201 DVKKREMYDMVGLSVNSYSDGLRIDTVKHYOKDFWPGYNKAAGYTCIGEVLDGDPAYTC 260
DB 182 SDVASVFNKDYKDFVGNYSIDGLRIDSAKHVDGFFPFVSAAGYVSGEVFGQDPAYTC 241

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:54:40 ; Search time 38.1 Seconds

(without alignments)
506.098 Million cell updates/sec

Title: US-09-710-339-2

Perfect score: 2684

Sequence: 1 MVAWMSIFLYGIQVAPALA.....LPRIYPTREKLAGSKICSSS 498

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2684	100.0	499	1	AMYA_ASPOR
2	2675	99.7	499	1	AMY ASPSH
3	2674	99.6	499	1	AMYB_ASPAW
4	2662	99.2	498	1	AMYA_ASPAW
5	1778	66.2	484	1	AMYA_ASPNG
6	1449	54.0	512	1	AMY1_DEBOS
7	1397.5	52.1	494	1	AMY1_SACFI
8	1015.5	37.8	513	1	AMY2_SCHPO
9	899	33.5	581	1	AMY2_SCHPO
10	898	33.5	478	1	YDDL_SCHPO
11	670	25.0	1196	1	AMYB_PABPO
12	456.5	17.0	713	1	CDGT_BACSO
13	454.5	16.9	712	1	CDGT_BACSO
14	444.5	16.6	713	1	CDGT_BACSP
15	440.5	16.4	528	1	AMY_BACCI
16	433	16.1	710	1	CDGT_THETU
17	428.5	15.9	655	1	CDGT_KLEPN
18	428	15.0	703	1	CDGT_BACS2
19	424.5	15.8	919	1	AMY_STRLI
20	418	15.6	714	1	CDGU_BACCI
21	418	15.6	713	1	CDGI_PAEWA
22	414.5	15.4	713	1	AMYR_BACS8
23	412	15.4	717	1	AMYM_BACST
24	407	15.2	704	1	CDGT_BACOH
25	406.5	15.1	718	1	CDGT_BACCI
26	405.5	15.1	718	1	CDGT_BACSS
27	405	15.1	713	1	CDG2_PAEWA
28	399.5	14.9	718	1	CDGT_BACLI
29	385.5	14.4	711	1	NEPU_BACST
30	368	13.7	588	1	NEPU_BACST
31	337	12.6	586	1	AMYM_BACAD
32	328.5	12.2	676	1	AMY1_ECCLT
33	317.5	11.8	591	1	CDAS_BACSH

34	315	11.7	562	1	AMY2_DICTH	P14898 dictyoglomu
35	295.5	11.0	574	1	CDAS_THETU	P29964 thermococcus
36	270	10.1	498	1	AMY2_DICTH	P14899 dictyoglomu
37	269	10.0	585	1	NEPU_THETU	Q08751 thermococcus
38	263	9.8	1481	1	APU_THETU	P38939 t amylopull
39	261	9.7	471	1	AMY_TENMO	P56634 tenebrio mo
40	259.5	9.7	518	1	AMT6_BACS7	P19571 bacillus sp
41	257	9.6	489	1	AMY_TRICA	P09107 tribolium c
42	255.5	9.5	561	1	TREC_BACSU	P39795 bacillus su
43	253.5	9.4	520	1	AMY_BACME	P20845 bacillus me
44	252.5	9.4	1475	1	APU_THETU	P16950 t amylopull
45	247	9.2	478	1	AMY_STRHY	P08486 streptomyce

ALIGNMENTS

```

RESULT 1
ID AMYA_ASPOR STANDARD: PRT; 499 AA.
AC P10529; P11763; Q00250;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amylase A precursor (EC 3.2.1.1) (Taka-amylase A) (TAA)
DE (1,4-alpha-D-glucan glucanohydrolase).
GN AMY1 AND AMY2 AND AMY3.
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 63303;
RX MEDLINE=89237897; PubMed=2785629;
RA Wirsel S., Lachmund A., Wildhardt G., Rutkowski E.;
RT "Three alpha-amylase genes of Aspergillus oryzae exhibit identical
RT intron-exon organization.";
RL Mol. Microbiol. 3:3-14(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89378767; PubMed=2789162;
RA Genes M.J., Dove M.J., Seliy V.L.;
RT "Aspergillus oryzae has two nearly identical Taka-amylase genes, each
RT containing eight introns.";
RL Gene 79:107-117(1989).
RN [3]
RP SEQUENCE FROM N.A.
RA Tada S., Imura Y., Gomi K., Takahashi K., Hara S., Yoshizawa K.;
RT "Cloning and nucleotide sequence of the genomic Taka-amylase A gene
RT of Aspergillus oryzae.";
RL Agric. Biol. Chem. 53:593-599(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=90128276; PubMed=2612911;
RA Tsukagoshi N., Furukawa M., Nagaba H., Kirita N., Tsuboi A.,
RA Ueda S.;
RT "Isolation of a cDNA encoding Aspergillus oryzae Taka-amylase A:
RT evidence for multiple related genes.";
RL Gene 84:319-327(1989).
RN [5]
RP SEQUENCE OF 22-499.
RA Toda H., Kondo K., Narita K.;
RT "The complete amino acid sequence of Taka-amylase A.";
RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 58:208-212(1982).
RN [6]
RP SEQUENCE OF 206-225.
RX MEDLINE=7401521; PubMed=4733850;
RA Isemura S., Ikenaka T.;
RT "The amino acid sequences of glycopeptides obtained from Taka-amylase
RT A with trypsin and chymotrypsin.";
RL J. Biochem. 74:1-10(1973).
RN [7]

```


Query Match 100.0%; Score 2684; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1, 1e-176;
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MWAMWSLFYGLQVAPALAAPADMRQSISYFLITDRFARTDGSSTATCMTADOKYCGG 60
DB 2 MWAMWSLFYGLQVAPALAAPADMRQSISYFLITDRFARTDGSSTATCMTADOKYCGG 61
OY 61 TWOGIIDKLDYIOGMGFTAITWPTAQLPQTAYGDAYHGYWODIYSLNENYGTADL 120
DB 62 TWOGIIDKLDYIOGMGFTAITWPTAQLPQTAYGDAYHGYWODIYSLNENYGTADL 121
OY 121 KALSALHERGMYLMDVYVANHMGYDAGSSVDYSVEKFPSSODYFHPFCFIQNYEDQTO 180
DB 122 KALSALHERGMYLMDVYVANHMGYDAGSSVDYSVEKFPSSODYFHPFCFIQNYEDQTO 181
OY 181 VEDCWLGNNTVSLPDLDTTKDYVKNEMVDWYGLSVNSYIDGLRIDYKHKVQKDFWPGYN 240
DB 182 VEDCWLGNNTVSLPDLDTTKDYVKNEMVDWYGLSVNSYIDGLRIDYKHKVQKDFWPGYN 241
OY 241 KAAGVYCIGEVLDGDPATCPYONMVGVLNPIYPLNFKSTSGSMDLYNMINTVK 300
DB 242 KAAGVYCIGEVLDGDPATCPYONMVGVLNPIYPLNFKSTSGSMDLYNMINTVK 301
OY 301 SDPCPSTLLGTFFVENHNDNPREFASYTNDIALAKNVAFFIINDGIPITTYAGOEHOHYAGND 360
DB 302 SDPCPSTLLGTFFVENHNDNPREFASYTNDIALAKNVAFFIINDGIPITTYAGOEHOHYAGND 361
OY 361 PANRATATLSCGPTDSELYKLASANAIRNVAISKDTGFVYTKKMPYKDDTTIAMRKCT 420
DB 362 PANRATATLSCGPTDSELYKLASANAIRNVAISKDTGFVYTKKMPYKDDTTIAMRKCT 421
OY 421 DGSQIVTITLSNKGASGDSYTLISGAGYTAGOOLTEVIGCTTIVTGSNGNVPVMAAGLP 480
DB 422 DGSQIVTITLSNKGASGDSYTLISGAGYTAGOOLTEVIGCTTIVTGSNGNVPVMAAGLP 481
OY 481 RVLYPTEKLAGSKICSS 498
DB 482 RVLYPTEKLAGSKICSS 499

```

RESULT 2
 ID AMY_ASPSH STANDARD: PRT: 499 AA.
 AC P30292;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
 DE glucanohydrolase).
 GN AMY.
 OS Aspergillus shirousami.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5070;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92323146; PubMed=1368777;
 RA Shibuya I., Tamura G., Ishikawa T., Hara S.;
 RT "Cloning of the alpha-amylase cDNA of Aspergillus shirousami and its
 RT expression in Saccharomyces cerevisiae.";
 RL Biosci. Biotechnol. Biochem. 56:174-179(1992).
 CC -1 CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -1 COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
 CC -1 SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----

```

DR EMBL; D10461; BAA01255.1; -
DR PIR; J50663; J50663.
DR HSSP; P10529; 77AA.
DR InterPro; IPR000461; Alpha-amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 499
FT ACT_SITE 227 227
FT ACT_SITE 231 231
FT ACT_SITE 251 251
FT ACT_SITE 318 318
FT ACT_SITE 51 59
FT DISULFID 171 185
FT DISULFID 261 304
FT DISULFID 461 496
FT CARBOHYD 218 218
SQ SEQUENCE 499 AA; 54852 MW; 1FB7AE50DA01C03F CRC64;

```

Query Match 99.7%; Score 2675; DB 1; Length 499;
 Best Local Similarity 99.8%; Pred. No. 4, 4e-176;
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 MWAMWSLFYGLQVAPALAAPADMRQSISYFLITDRFARTDGSSTATCMTADOKYCGG 60
DB 2 MWAMWSLFYGLQVAPALAAPADMRQSISYFLITDRFARTDGSSTATCMTADOKYCGG 61
OY 61 TWOGIIDKLDYIOGMGFTAITWPTAQLPQTAYGDAYHGYWODIYSLNENYGTADL 120
DB 62 TWOGIIDKLDYIOGMGFTAITWPTAQLPQTAYGDAYHGYWODIYSLNENYGTADL 121
OY 121 KALSALHERGMYLMDVYVANHMGYDAGSSVDYSVEKFPSSODYFHPFCFIQNYEDQTO 180
DB 122 KALSALHERGMYLMDVYVANHMGYDAGSSVDYSVEKFPSSODYFHPFCFIQNYEDQTO 181
OY 181 VEDCWLGNNTVSLPDLDTTKDYVKNEMVDWYGLSVNSYIDGLRIDYKHKVQKDFWPGYN 240
DB 182 VEDCWLGNNTVSLPDLDTTKDYVKNEMVDWYGLSVNSYIDGLRIDYKHKVQKDFWPGYN 241
OY 241 KAAGVYCIGEVLDGDPATCPYONMVGVLNPIYPLNFKSTSGSMDLYNMINTVK 300
DB 242 KAAGVYCIGEVLDGDPATCPYONMVGVLNPIYPLNFKSTSGSMDLYNMINTVK 301
OY 301 SDPCPSTLLGTFFVENHNDNPREFASYTNDIALAKNVAFFIINDGIPITTYAGOEHOHYAGND 360
DB 302 SDPCPSTLLGTFFVENHNDNPREFASYTNDIALAKNVAFFIINDGIPITTYAGOEHOHYAGND 361
OY 361 PANRATATLSCGPTDSELYKLASANAIRNVAISKDTGFVYTKKMPYKDDTTIAMRKCT 420
DB 362 PANRATATLSCGPTDSELYKLASANAIRNVAISKDTGFVYTKKMPYKDDTTIAMRKCT 421
OY 421 DGSQIVTITLSNKGASGDSYTLISGAGYTAGOOLTEVIGCTTIVTGSNGNVPVMAAGLP 480
DB 422 DGSQIVTITLSNKGASGDSYTLISGAGYTAGOOLTEVIGCTTIVTGSNGNVPVMAAGLP 481
OY 481 RVLYPTEKLAGSKICSS 498
DB 482 RVLYPTEKLAGSKICSS 499

```

RESULT 3
 ID AMYB_ASPAW STANDARD: PRT: 499 AA.
 AC 002906;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Alpha-amylase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase B).
CN AMYB.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OVK143F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
RT from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52756; CAA36967.1; -.
DR HSSP: P10529; TTA.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 499 ALPHA-AMYLASE B.
FT ACT_SITE 227 227 BY SIMILARITY.
FT ACT_SITE 231 231 BY SIMILARITY.
FT ACT_SITE 251 251 BY SIMILARITY.
FT ACT_SITE 318 318 BY SIMILARITY.
FT DISULFID 51 59 BY SIMILARITY.
FT DISULFID 171 185 BY SIMILARITY.
FT DISULFID 261 304 BY SIMILARITY.
FT DISULFID 461 496 BY SIMILARITY.
FT CARBOHYD 218 218 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 499 AA; 54921 MW; 740B96B11BC01A8A CRC64;
Query Match 99.6%; Score 2674; DB 1; Length 499;
Best Local Similarity 99.6%; Pred. No. 5.2e-176;
Matches 496; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 MWAMMSFLYGLGVAPAPALAAIPADWRSQSIYFLITDFARTDSTATATCTADOKYCGG 60
DB 2 MWAMMSFLYGLGVAPAPALAAIPADWRSQSIYFLITDFARTDSTATATCTADOKYCGG 61
OY 61 TWGIIIDKLDYIQGMGFATWITPTAOLPOTAYGDAHGYWOODIYSLNENGTADDL 120
DB 62 TWGIIIDKLDYIQGMGFATWITPTAOLPOTAYGDAHGYWOODIYSLNENGTADDL 121
OY 121 KALSSALHBRGMYLWVDVYANMGYDAGSSVDYVFPFSSODYFHFCTIQNEDOTQ 180
DB 122 KALSSALHBRGMYLWVDVYANMGYDAGSSVDYVFPFSSODYFHFCTIQNEDOTQ 181
OY 181 VEDCWLGNTVYSLPDLDTTKDYVKNENYDWGSLVSNYSIDGLRIDYKHYQKDFWPGYN 240
DB 182 VEDCWLGNTVYSLPDLDTTKDYVKNENYDWGSLVSNYSIDGLRIDYKHYQKDFWPGYN 241
OY 241 KAAGVYCIGEVLDGDPAYTCYQNVMDGVLPYIYPLLNAFKSTSGMDLYNMINTVK 300
|||||

DB 242 KAAGVYCIGEVLDGDPAYTCYQNVMDGVLPYIYPLLNAFKSTSGMDLYNMINTVK 301
OY 301 SDCPDSTLIGTVEVNDHNDPRFASYNTNDIALKNVAFITLNDGJPIIYAGGEOHYAGND 360
DB 302 SDCPDSTLIGTVEVNDHNDPRFASYNTNDIALKNVAFITLNDGJPIIYAGGEOHYAGND 361
OY 361 PANREATWLSGYPTDSELYKLJASANAIRNVAISKDTGEVYTKMNPYKDDTTIAMRGT 420
DB 362 PANREATWLSGYPTDSELYKLJASANAIRNVAISKDTGEVYTKMNPYKDDTTIAMRGT 421
OY 421 DGSQVITLISNKGASDSTTSLSGAGCTAGGQULTEVIGCTTYVYGSQNVVPWAGSLP 480
DB 422 DGSQVITLISNKGASDSTTSLSGAGCTAGGQULTEVIGCTTYVYGSQNVVPWAGSLP 481
OY 481 RVLYPTEKLKSGKICSS 498
DB 482 RVLYPTEKLKSGKICSS 499
RESULT 4
ID AMYA.ASPAW STANDARD; PRT; 498 AA.
AC 002905;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Alpha-amylase A precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase A).
GN AMYA.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OVK143F;
RX MEDLINE=90254827; PubMed=2340591;
RA Korman D.R., Bayliss F.T., Barnett C.C., Carmona C.L.,
RA Kodama K.H., Royer T.J., Thompson S.A., Ward M., Wilson L.J.,
RA Berka R.M.;
RT "Cloning, characterization, and expression of two alpha-amylase genes
RT from Aspergillus niger var. awamori.";
RL Curr. Genet. 17:203-212(1990).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52755; CAA36966.1; -.
DR HSSP: P10529; TTA.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 498 ALPHA-AMYLASE A.
FT ACT_SITE 227 227 BY SIMILARITY.
FT ACT_SITE 231 231 BY SIMILARITY.
FT ACT_SITE 251 251 BY SIMILARITY.
FT ACT_SITE 318 318 BY SIMILARITY.
FT DISULFID 51 59 BY SIMILARITY.
FT DISULFID 171 185 BY SIMILARITY.
FT DISULFID 261 304 BY SIMILARITY.
FT DISULFID 461 496 BY SIMILARITY.

FT CARBOHYD 218 218 N-LINKED (GLCMNC. . .) (POTENTIAL).
SQ SEQUENCE 498 AA; 54880 MM; 7658511BC01A8A01 CRC64;

Query Match 99.2%; Score 2662; DB 1; Length 498;
Best Local Similarity 99.6%; Pred. No. 3.4e-175;
Matches 493; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MYAMMSLEFLGQVAPALATPADMRSOSIYELTDREARTDGTATCMTADOKYCGG 60
DB 2 MYAMMSLEFLGQVAPALATPADMRSOSIYELTDREARTDGTATCMTADOKYCGG 61
QY 61 TWOGIDKLDYIOGMGTATWIPVTAQLPOTATAGAHYGYOQDIYSLNENYGTADLL 120
DB 62 TWOGIDKLDYIOGMGTATWIPVTAQLPOTATAGAHYGYOQDIYSLNENYGTADLL 121
QY 121 KALSALHERGMYLMDVYVANHMGYDAGSSVDYSVFKPSSODYFHPFCIONYEDQTO 180
DB 122 KALSALHERGMYLMDVYVANHMGYDAGSSVDYSVFKPSSODYFHPFCIONYEDQTO 181
QY 181 VEDCMLGDNVYSLPDLDTTKDYVANKENYDMVGLSVSNYSIDGLRIDTVKHKQDFWPGTN 240
DB 182 VEDCMLGDNVYSLPDLDTTKDYVANKENYDMVGLSVSNYSIDGLRIDTVKHKQDFWPGTN 241
QY 241 KAAGYTCIGEVLDGDPATYCPYQVMDGVNLYPYPLNAFKSTSGMDLYNMTNFK 300
DB 242 KAAGYTCIGEVLDGDPATYCPYQVMDGVNLYPYPLNAFKSTSGMDLYNMTNFK 301
QY 301 SDPCDSTLLGTFEVENHNDPRFASVTNDIALAKNVAFFIILNDGPIIYAGOEQHYAGND 360
DB 302 SDPCDSTLLGTFEVENHNDPRFASVTNDIALAKNVAFFIILNDGPIIYAGOEQHYAGND 361
QY 361 PANRATATLSCYPTDSELYKLIASNAIRNVAISKOTGFYTKNMPYKDDTTIARKEGI 420
DB 362 PANRATATLSCYPTDSELYKLIASNAIRNVAISKOTGFYTKNMPYKDDTTIARKEGI 421
QY 421 DSGQIVTILSNKKGASGDSYTLSELGAGYTAGOQLTEYIGCTTVYVSGDGNVPYPMAGLP 480
DB 422 DSGQIVTILSNKKGASGDSYTLSELGAGYTAGOQLTEYIGCTTVYVSGDGNVPYPMAGLP 481
QY 481 RVLVPTKELAGSKTC 495
DB 482 RVLVPTKELAGSKTC 496

RESULT 5

AMYA_ASPNG STANDARD; PRT: 484 AA.
AC P36271.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Acid alpha-amylase (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eutotiomyces;
OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN (1)
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=91002514; PubMed=2207069;
RA Boel E., Brady L., Brzozowski A.M., Derewenda Z., Dodson G.G.,
Jensen V.J., Petersen S.B., Swift H., Thim L., Woldike H.F.,
"Calcium binding in alpha-amylases: an X-ray diffraction study at
2.1-A resolution of two enzymes from Aspergillus";
RT Biochemistry 29:6244-6249(1990).
CC -I- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
linkages in oligosaccharides and polysaccharides.
CC -I- COPACOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR PDB: 2AAA; 15-JUL-93.
DR InterPro: IPR00461; Alpha_amylase.
DR Pfam: PF00128; alpha-amylase; 1.

KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; 3D-structure.

FT ACT_SITE 206 206
FT ACT_SITE 210 210
FT ACT_SITE 230 230
FT ACT_SITE 297 297
FT ACT_SITE 30 30
FT DISULFID 150 164
FT DISULFID 240 283
FT DISULFID 440 475
FT CARBOHYD 24 24
FT CARBOHYD 157 157
FT CARBOHYD 197 197
SQ SEQUENCE 484 AA; 52935 MM; 04D596B3468056D CRC64;

Query Match 66.2%; Score 1778; DB 1; Length 484;
Best Local Similarity 66.9%; Pred. No. 1.2e-114;
Matches 318; Conservative 63; Mismatches 94; Indels 0; Gaps 0;

QY 24 ADMRSOSIYELTDREARTDGTATCMTADOKYCGGTWQGITDKLDYIOGMGTATWIP 83
DB 4 ASMRQOSIYELTDREARTDGTATCMTADOKYCGGTWQGITDKLDYIOGMGTATWIP 83
QY 84 PVTAQLPOTATAGAHYGYOQDIYSLNENYGTADLLKALSALHERGMYLMDVYVANH 143
DB 64 PITDOLPOTATAGAHYGYOQDIYSLNENYGTADLLKALSALHERGMYLMDVYVANH 123
QY 144 GYDAGSSVDYSVFKPSSODYFHPFCIONYEDQTOVEDCMLGDNVYSLPDLDTTKDY 203
DB 124 GYDAGSSVDYSVFKPSSODYFHPFCIONYEDQTOVEDCMLGDNVYSLPDLDTTKDY 183
QY 204 KNEWDMWGLSVSNYSIDGLRIDTVKHKQDFWPGTNKAAGYTCIGEVLDGDPATYCPY 263
DB 184 KNEWDMWGLSVSNYSIDGLRIDTVKHKQDFWPGTNKAAGYTCIGEVLDGDPATYCPY 243
QY 264 NVMDGVNLYPYPLNAFKSTSGMDLYNMTNFKSDPCDSTLLGTFEVENHNDPRFAS 323
DB 244 KVLGDLVNLVYPLWQLDAFESSSSISNLNMTIKSVASDCDPTLLGNFLENHNDPRFAS 303
QY 324 YTNDAALAKNVAFFIILNDGPIIYAGOEQHYAGNDPANRATATLSCYPTDSELYKLI 383
DB 304 YTSDSQAKNVAFFIILNDGPIIYAGOEQHYAGNDPANRATATLSCYPTDSELYKLI 363
QY 384 SANAIRNVAISKOTGFYTKNMPYKDDTTIARKEGDSQIVTILSNKKGASGDSYTLSEL 443
DB 364 TTNAIRKLIADASAYITVANDAFYTDNSNTIARKEGDSQIVTILSNKKGASGDSYTLSEL 423
QY 444 SGAGYTAGOQLTEYIGCTTVYVSGDGNVPYPMAGLPYKDDTTIARKEGDSYTLSEL 498
DB 424 SGAGYTAGOQLTEYIGCTTVYVSGDGNVPYPMAGLPYKDDTTIARKEGDSYTLSEL 478

RESULT 6

AMY1_DEBOC STANDARD; PRT: 512 AA.
AC P19269;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amylase 1 precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
glucanohydrolase).
GN AMY1.
OS Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=27300;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 26076;
RX MEDLINE=90032659; PubMed=2806251;
RA Strasser A.W.M., Selk R., Dohmen R.J., Niemann T., Bielefeld M.,
Seebach P., Tu G., Hollenberg C.P.;

RT "Analysis of the alpha-amylase gene of Schwannomyces occidentalis
RT and the secretion of its gene product in transformants of different
RT yeast genera."
RT Eur. J. Biochem. 184:699-706(1989).
RT [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CCRC 21164;
RX MEDLINE-92120467; PubMed-1769525;
RA Wu F.M., Wang T.T., Hsu W.H.;
RT "The nucleotide sequence of Schwannomyces occidentalis alpha-amylase
RT gene."
RT FEMS Microbiol. Lett. 66:313-318(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 26077 / CBS 2863;
RX MEDLINE-92307400; PubMed-1612414;
RA Park J.C., Bai S., Tai C.Y., Chun S.B.;
RT "Nucleotide sequence of the extracellular alpha-amylase gene in the
RT yeast Schwannomyces occidentalis ATCC 26077."
RT FEMS Microbiol. Lett. 72:17-23(1992).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides
CC -1- COPROCTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- ENZYME REGULATION: ALPHA-AMYLASE EXPRESSION UNDERLIES CATABOLITE
CC REPRESSION BY GLUCOSE.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S77586; AAB21151.2; -
DR EMBL: X16040; CAA34162.1; -
DR EMBL: X62079; CAA43995.1; -
DR EMBL: S38381; AAB22383.2; -
DR PIR: S06115; S06115.
DR PIR: S23355; S23355.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Calcium;
KW Glycoprotein; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 512 ALPHA-AMYLASE 1.
FT ACT_SITE 242 242 BY SIMILARITY.
FT ACT_SITE 246 246 BY SIMILARITY.
FT ACT_SITE 333 333 BY SIMILARITY.
FT CARBOHYD 233 233 N-LINKED (GLUCAN. . .) (PROBABLE).
FT DISULFID 66 74 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 276 319 BY SIMILARITY.
FT DISULFID 475 510 BY SIMILARITY.
FT VARIANT 32 32 M->K (IN STRAINS CCRC 21164 AND ATCC
FT 26077).
FT VARIANT 36 36 S->G (IN STRAIN CCRC 21164).
FT VARIANT 73 73 Y->I (IN STRAIN ATCC 26077).
FT VARIANT 280 280 N->S (IN STRAIN CCRC 21164).
FT VARIANT 350 350 D->A (IN STRAINS CCRC 21164 AND ATCC
FT 26077).
FT VARIANT 479 479 L->S (IN STRAINS CCRC 21164 AND ATCC
FT 26077).
FT VARIANT 483 483 S->F (IN STRAIN CCRC 21164).
FT SEQUENCE 512 AA; 56527 MW; 857552B2C60F965 CRC64;
Query Match 54.0%; Score 1449; DB 1; Length 512;
Best Local Similarity 56.4%; Pred. No. 4.4e-92;
Matches 269; Conservative 74; Mismatches 132; Indels 2; Gaps 2;

OY 21 AAPAD-WRSQSTIELLTDREARFDGTTATCMTADQKRCGCTWOGIIDKLDYIOGKCF7A 79
OY 36 SSADPKMKRQSIYQIVTRFARSDGSTADCLVSDRYCKSGYKGIIDKLDYIOGKCF7A 95
OY 80 IWTIVTQLOPQTATYAGVAGHYGMOOIVSLNENYGRADDLKALSSALHBRGYLWADV 139
OY 96 IWTISVEQIPONTATYGAIGHYMKNDLNTNFGTADDELKOLASLHRSKMLWADV 155
OY 140 ANHMGYDAGSSVDYVSKPPSSODYFPFCFIONTEYEDQIOWEDCWGLDNTVSLPDLDTT 199
OY 156 YNHYAMNGDGSVDYSPFPNOOSYFHDYCLITNYNDQFVWEDCWGDETEVSLPDLSTE 215
OY 200 KDVKNENWMDVGSIVSNYSIDGLRIDYKHYQDQFMPGVNKAAGYCIETVSDSPATV 259
OY 216 DNEVIGFQJWSDVFNQYSDIDGLRIDSAKHYDTASLTKEEDAGVNLDEYVQGGDTT 275
OY 260 CPYQNVMDGLNYPYPLBLNAFKSTSGSMDLYNMINTVKSDDPSTLLGTVEVHNHP 319
OY 276 CPYQNVMDGLNYPYPLBLNAFKSTSGSMDLYNMINTVKSDDPSTLLGTVEVHNHP 335
OY 320 RASVTNDIADLAKNVAELIINDGPIYAGORQHYAGNDPNAFRTMTLSGYPPTSELY 379
OY 336 RPPSVTSPTSLIKNMAFIILADGPIYIYQEQGLNGSDPARRKALMLSGINTDSEY 395
OY 380 KLIASANAIKNYAIKSDYFYTKNMPYKDDTTIARKKGTDSQIVTILNKGAGSDSY 439
OY 396 ELISKLNQIRNOAIKKSASVYKSSVSSDHYIATRKSGDANQILISFNILGNSNG-SQ 454
OY 440 TISLGAGTAAQOLTEYIGCTVYVGSQVPPMAGCLPRVLYPEKLAGSKYS 496
OY 455 DTVSNVTGSSDVKYDIIISCNVLAGDSLSVSISSGMPVYAFPSVLSGSDCN 511
RESULT 7
ID AMY1_SACFI STANDARD; PRT; 494 AA.
AC P21567;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Alpha-amylase precursor (EC 3.2.1.1).
GN ALPI.
OS Saccharomycopsis fibuligera (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycopsidaceae; Saccharomycopsis.
OX NCBI_TaxID=4944;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87276512; PubMed-3497057;
RA Itoh T., Yamashita I., Fukui S.;
RT "Nucleotide sequence of the alpha-amylase gene (ALPI) in the yeast
RT Saccharomycopsis fibuligera."
RT FEBS Lett. 219:339-342(1987).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COPROCTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X05791; CAA29233.1; -
DR PIR: S00064; ALBYAF.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha amylase.
DR Pfam: PF00128; alpha-amylase; 1.

```

RP SEQUENCE FROM N.A.
RC STRAIN=972:
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.:
RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Endohydrolays of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z98762; CAB1471.1; -.
CC DR HSPSP; P10529; 77PA.
CC DR InterPro: IPR00461; Alpha-amylase.
CC DR Pfam: PF00128; alpha-amylase; 1.
CC KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
CC Calcium; Glycoprotein; Signal.
CC KW SIGNAL 1 25 POTENTIAL.
CC FT CHAIN 26 513 PROBABLE ALPHA-AMYLASE C4A8.01.
CC FT ACT_SITE 226 226 BY SIMILARITY.
CC FT ACT_SITE 230 230 BY SIMILARITY.
CC FT ACT_SITE 250 250 BY SIMILARITY.
CC FT ACT_SITE 318 318 BY SIMILARITY.
CC FT DISULFID 52 60 BY SIMILARITY.
CC FT DISULFID 171 184 BY SIMILARITY.
CC FT DISULFID 260 304 BY SIMILARITY.
CC FT DISULFID 454 488 BY SIMILARITY.
CC FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 513 AA; 58715 MW; 455DD97FA428C182 CRC64;

Query Match 37.8%; Score 1015.5; DB 1; Length 513;
Best local Similarity 41.4%; Pred. No. 2,2e-62;
Matches 204; Conservative 84; Mismatches 170; Indels 35; Gaps 12;

QY 5 WSL-FLYLQVLAARALATPADWRSQSYFLTLTRFARTDSTATCNTADQKCGSTWQ 63
DB 9 WTLKAVSLFLVATIPNALDKHGMRKQSYSLTLTRFASTN--KPCNPEDREYCGGNMR 65
QY 64 GIIDKLADYIQMGFTAIWITPVTALPQTVYGDVAGYQMOODIYSINENYGTADDLKAL 123
DB 66 GIIDKLADYIQMGFTAIWITPVTALPQTVYGDVAGYQMOODIYSINENYGTADDLKAL 125
QY 124 SSALHERGMTLVADVANHMGYDAGSSVDYSVKKPFSSQDYFHFPCFIQNYEDQTVED 183
DB 126 ADALHDERMTLVADVANHMG-SSDPKNIDYGIKPPNOSHHPMCPRIQ-DKPLSLFQ 183
QY 184 CMLGDNRTVSLDDLTPTDYVKNEMYDWVGSLSYKSIDGRIQVKNVQKDFMPGYKKA 243
DB 184 CMIGTEDEMTLPDIITENPQIIETLYNTHDQVQKFKIDGLAVDQTKVTRRTFWMGFCESA 243
QY 244 GVCYIGEVLDGDPAYTQCYQNVMDGVNLYPI-----YYPLIN-AFKSTSGSDMLYNN 295
DB 244 GVCYCGEEMTGOADLFCEWQDEYMDGLHNFPQGVAAESVIPLNDRLARKTAIAM----- 297
QY 296 INTVKSDCPSTLTGTVENHDNRPASFYINDIALANVAAFITLNDGPIITYAGQEQHY 355
DB 298 -NLVAHHCKOSTLTGLTLESODAPRIALALNNDDYVILNMAFTLNMSSGPIIVYGOEQMP 356
QY 356 AGNDPDPANREATWLSGPTDSELYKFLASANAIRNYAISKDTG--FTYKKNWPIYKDDTT 413
DB 357 NGSHDVPNRRPALMDQGVNTDGPITQYISKVKKIKRDLINSEDEIYIRSTTHAIMGDHY 416
QY 414 TAMKCRGDSQTVYITENKCRASQDSQVATGCRAGYAKQOQITAEVIGSTRTVTCQNTY 416

```

```

Db 417 MWMTKG-----PVITFTNYGAVDKEYLIKMPGS-----EMMIDLTCTLIEV--EGEVMR 465
QY 473 VPMAGLPRVLYP 485
Db 466 TSIRKGEKRLYP 478

RESULT 9
AMY1_SCHPO STANDARD; PRT; 581 AA.
ID AMY1_SCHPO
AC 009840:
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE Probable alpha-amylase C23D3.14C precursor (EC 3.2.1.1) (1,4-alpha-D-
DE glucan glucanohydrolase).
GN SPAC23D3.14C.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Nidlett D., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 264354; CAA91249.1; -.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR00461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
KW Calcium; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 581 PROBABLE ALPHA-AMYLASE C23D3.14C.
FT ACT_SITE 233 233 BY SIMILARITY.
FT ACT_SITE 237 237 BY SIMILARITY.
FT ACT_SITE 325 325 BY SIMILARITY.
FT DISULFID 56 64 BY SIMILARITY.
FT DISULFID 176 191 BY SIMILARITY.
FT DISULFID 267 311 BY SIMILARITY.
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 581 AA; 67004 MW; FE9DE9D323E1890 CRC64;

Query Match 33.5%; Score 899; DB 1; Length 581;
Best Local Similarity 36.2%; Pred. No. 2.4e-54;
Matches 180; Conservative 94; Mismatches 207; Indels 16; Gaps 8;
QY 2 VAWMSIFELXGLOVAAAPALATPA-DMRSOSIYELLTRFRARDGST-TATCNTADQKYCG 59
Db 11 IGMMLLFAF-----IPAYAGHSAEMKRSRISYQITDRSLEGAETRLPCDPVPMFCG 65
QY 60 GTWQGIIDKLDYIQGKGFALMTPTTAQLPQTAYGDAHYHGYWODIYSLNENGTAD 119
Db 66 GTWNGIRNHLIDYIOGKGFDAIMTSPITFENVECDNDIDSSYHWTNTNLIELNHGTRKEE 125

```

```

QY 120 LKALSALHERGYLMVDVYANHMGVDGASSVDYVFEPFSSODYFHPFCEIONYE--D 177
Db 126 FMEILIOELHKKRDWILDLDAINSMALINGPLDQMSEKVIYFPDASFPHRCWV-DYESND 184
QY 178 QTVQEDCWLGMDTVSLPDDITTKDVKYKNEMYDVGSLVSNYSIDGIRIDYKHKVOKDFWP 237
Db 185 IBSVQNCWLGDEBNLLADVDTENEVVLVLEKMKVNVQVEYDIDGIRFDAIKHAFIEFWL 244
QY 238 GYNKAAGVYCIGEVLDGDPAITCPYQNV-MDGLVNPPIYPLLMFKPSSGMDLYMI 296
Db 245 RSKKADIFTTIGETFTGSPAEACDYNSGLDSEFLNPEPLTWMFNMNGLOCCELAIAI 304
QY 297 NTFKSDCPDSTLLGTFVEHNDNPRFASYNTDIALAKNVAFTIINDGPIIYAGOECHYA 356
Db 305 NQINECDNDINVIQGTIGNHDLPRISHNNTDQARIMNATTFVMMMDGPIIYGTGEQNFN 364
QY 357 GGNDPANREATWLSGYPTDSELYKLIAANAIKNAISKDTGEPVYKKMPTIKDDTTAM 416
Db 365 SYHDPFNREALMNSNDMENRYKKLIGILNRFKRSVQROEYVNTNSTILSVKIHHTIV 424
QY 417 RKGTDSQIVTILSNKGASGDS-YTSLSGAGTACQOLTEVYIGCTTVYVGSQNVYPM 475
Db 425 QK---LVNITVLNNGIHNERLSIVKPLGASPKOTFTIINOKYVNTDGLKAYI 480
QY 476 AGGLPRVLYPTEKLAGS 492
Db 481 TNGFPVLYPTSKIRTS 497

RESULT 10
YDYL_SCHPO STANDARD; PRT; 478 AA.
ID YDYL_SCHPO
AC 010427; OGUUN0.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 55.4 kDa protein C188.01c in chromosome III.
GN SPC188.01C OR SPC11E10.09C.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Ramsperger U., Pohl T., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL121783; CAB57851.1; -.
DR HSSP: AL049662; CAB41221.1; -.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR00461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KW Hypothetical protein; Hydrolase; Glycosidase.
FT ACT_SITE 216 216 BY SIMILARITY.
FT ACT_SITE 220 220 BY SIMILARITY.
FT ACT_SITE 308 308 BY SIMILARITY.
FT SEQUENCE 478 AA; 55442 MW; 3C65EAD005DF73E CRC64;

Query Match 33.5%; Score 898; DB 1; Length 478;

```


FT CONFLICT 736 736 N -> K (IN REF. 3).
 FT CONFLICT 741 741 N -> S (IN REF. 3).
 FT CONFLICT 758 758 S -> N (IN REF. 3).
 SQ SEQUENCE 1196 AA: 130893 MW: A41EA6B70F257064 CRC64:

Query Match 25.0%; Score 670; DB 1; Length 1196;
 Best Local Similarity 34.2%; Pred. No. 3e-38;
 Matches 172; Conservative 82; Mismatches 177; Indels 72; Gaps 17;

QY 13 QVAPALAAAP-----ADMRSGSIYFILDRFARPDGSGT---ATCTWADO 55
 DB 718 QTEPSPGAFISMGLTSLNPTSLMTDMSKOSITIFIMDRSNGDPSNDYGCNSNSDQ 777
 QY 56 -KYCGWQODIYSLNENYGTADIKALSSALHERGMATLVYVANHMGIDAGSSVDYSVKRPSSODYFHFCTI-- 114
 DB 778 RKMHGDPGQGIINKLDYIKMNGFAIMITPVMQKSEY-----AYHGHTYDFEAVDGH 832
 QY 115 GTADDLKALSSALHERGMATLVYVANHMGIDAGSSVDYSVKRPSSODYFHFCTI-- 172
 DB 833 GTMDKLELYRKAKHAKNIAYVAVVNHGDPFGNGFARA---PEFKADWYHHNGDITD 889
 QY 173 --QNYEOTQVEDCWLGDNTSLPDLDTKDYANKENYMDWVGSIVSNYSIDGLRIDTVKH 230
 DB 890 GDYNSNMWKIEN---GD-VAGLDDLHNEPNATNELKNMIMKMLNETGIDGLRIDTVKH 945
 QY 221 VQKDFWPGYKNRAGVYICGEVLDDPAVTCYQNVMDGVNLPIYPLNLNF-----K 283
 DB 946 VPKGFLKQFQDAANTFTMGELFHFHGDPAVYGDYTRYRLDAALDFPMYVYIKVFGHDSMRK 1005
 QY 284 STSGMDLYMIMTVKSDCDSLLTGFEVHNHNPFR-----ASYTNDIALAKNV 334
 DB 1006 IKDRYSDDRYR-----DAQTNGVFDNHVYKRFELANDASKPGANNDKMQOL-KAA 1055
 QY 335 AAFILNDGIPILIYAGOEYHAGNDPANREATWLSGYPDSELYKLASANAIRNYAIS 394
 DB 1056 LGFTLISGIPILITVGGTGGYSGGDDPANREN---MNFNAHNDLXYQYIAKINYRNHHPA 1112
 QY 395 KDTGCVYTKNMPYKDDPTIAMRKGTGSSQIVTILSKNGAGSDSYTSLSG-AGYTAGQO 453
 DB 1113 LQNGSQREK---WVDSFYSFQSKNGDEAIVFLNN---SWNSQTRITGNFMDLSMCTR 1165
 QY 454 LFEVIGCTVTVGSDGNVPVMA 476
 DB 1166 LFNOLNSDSVQI-NGSITVILA 1187

RESULT 12
 CDGT_BACSO STANDARD: PRT: 713 AA.
 ID CDGT_BACSO STANDARD: PRT: 713 AA.
 AC P05618:
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Cyclomaltodextrin glucanotransferase precursor (EC 2.4.1.19)
 DE (Cyclodextrin-glycosyltransferase) (CGTase).
 GN CGT.
 OS Bacillus sp. (strain 1011).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1410;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87308036; PubMed=2957361;
 RA Kimura K., Kataoka S., Ishii Y., Takano T., Yamane K.;
 RT "Nucleotide sequence of the beta-cyclodextrin glucanotransferase gene
 of alkalophilic Bacillus sp. strain 1011 and similarity of its amino
 acid sequence to those of alpha-amylases.";
 RL J. Bacteriol. 169:4399-4402(1987).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 RA Harata K., Haga K., Nakamura A., Aoyagi M., Yamane K.;
 RT "X-ray structure of cyclodextrin glucanotransferase from alkalophilic

RT Bacillus Sp. 1011. Comparison of two independent molecules at 1.8-A
 resolution.";
 RL Acta Crystallogr. D 52:1136-1145(1996).
 CC -1- CATALYTIC ACTIVITY: Degrades starch to cyclodextrins by formation
 of a 1,4-alpha-D-glucosidic bond.
 CC -1- COFACTOR: BINDS TWO CALCIUM IONS.
 CC -1- SUBUNIT: MONOMER.
 CC -1- MISCELLANEOUS: CGTASE MAY CONSIST OF TWO PROTEIN DOMAINS: THE ONE
 IN THE AMINO-TERMINAL SIDE CLEAVES THE ALPHA-1,4-GLUCOSIDIC BOND
 IN STARCH, AND THE OTHER IN THE C-TERMINAL SIDE CATALYZES OTHER
 ACTIVITIES, INCLUDING THE RECONSTITUTION OF AN
 ALPHA-1,4-GLUCOSIDIC LINKAGE FOR CYCLIZING THE
 MALTOLOGSACCHARIDE PRODUCED.
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M17366; AAA22308.1; .
 DR PIR: A26678; ALBSG1.
 DR PDB: 1PAM; 11-JAN-97.
 DR InterPro: IPR002044; Alpha_amylase.
 DR InterPro: IPR002909; IPT_TIG.
 DR pfam: PF00128; alpha-amylase; 1.
 DR pfam: PF02806; alpha-amylase_C; 1.
 DR pfam: PF00686; CBD_4; 1.
 DR pfam: PF01833; TIG; 1.
 DR PRODOM: PD001568; CBD_4; 1.
 DR Transfaserase; Glycosyltransferase; Calcium; Signal; 3D-structure.
 KW TRANSFERASE; 1
 FT CHAIN 28 713
 FT DOMAIN 28 165 A1.
 FT DOMAIN 166 229 B.
 FT DOMAIN 230 433 A2.
 FT DOMAIN 434 522 C.
 FT DOMAIN 523 609 D.
 FT DOMAIN 610 713 E.
 FT DISULFID 70 77
 FT ACT_SITE 256 256 BY SIMILARITY.
 FT ACT_SITE 284 284 BY SIMILARITY.
 FT ACT_SITE 355 355 BY SIMILARITY.
 SQ SEQUENCE 713 AA: 78340 MW: 524B259526B56C52 CRC64:

Query Match 17.0%; Score 456.5; DB 1; Length 713;
 Best Local Similarity 28.3%; Pred. No. 6.5e-24;
 Matches 149; Conservative 73; Mismatches 216; Indels 89; Gaps 21;

QY 1 MVAWSLFL-YGQVAPALAA-----TPADMRSGSIYFILDRF-----ART 42
 DB 7 LTAWTLMLELTGLISPAHAAPDTSVNKQNSTVDYITQIFDRSDGNPANNPAGAF 66
 QY 43 DGSFTATCNTADOKYCGWQGIIDKLD--YIOGFTAITWITPVAOLPQTTAY-----G 96
 DB 67 DGSCITNL-----RLYCGGDMQGIINKINDGYLNGMGITAIWISQPEVNTSVYNSGVNN 121
 QY 97 DAYHGWQODIYSLNENYGTADIKALSSALHERGMATLVYVANHMGIDAGSSVDYSVKRPSSODYFHFCTI-- 144
 DB 122 TAYHGYWADFCKKNPAPYGTMODFKNLIDTAHNNHNIIVIDFAPNHTSPASSDDPSFAEN 181
 QY 145 ---YDAGSSVDYSVKRPSSODYFHF-----PCEFTON--YEDOTQVEDCWLGDNTVSLP 194
 DB 182 GRLYDNGNLGGYT---NDQNLFFHHYGGTDSITENGILYKLYDLADLNHNSSVDY- 236
 QY 195 DLDITKDYVKNENYMDWVGSIVSNYSIDGLRIDTVKHV---QKDFWPGYKNRAGVYICGE 250

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 13, 2002, 08:53:25 ; Search time 158.68 Seconds
(without alignments)
7.700 Million cell updates/sec

Title: US-09-710-339-2_COPY_150_160
Perfect score: 57
Sequence: 1 SSVSDYSVFKEPF 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

arched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	478	15 AAR46065	Mutant alpha-amy
2	57	100.0	478	16 AAR79025	Mature taka-amy
3	57	100.0	478	16 AAR72450	Aspergillus oryzae
4	57	100.0	478	16 AAR78270	Aspergillus oryzae
5	57	100.0	478	17 AAR14500	Aspergillus oryzae
6	57	100.0	498	22 AAB84206	Amino acid sequenc
7	45	78.9	511	9 AAP81161	Recombinant alpha-
8	45	78.9	511	9 AAP81180	Sequence of alpha-
9	45	78.9	512	11 AAR07574	Alpha-amyase emco
10	41	71.9	493	17 AAR88212	Alpha-amyase. Th
11	39	68.4	5250	21 AAY53678	Sequence g1/442661

12	39	68.4	5322	22 ABB69422	Drosophila melanog
13	38	66.7	55	20 AAY41363	Human secreted pro
14	38	66.7	413	21 AAB10970	H. tuberculata hem
15	38	66.7	413	21 AAB71082	H. tuberculata Hth
16	38	66.7	413	22 AAB71137	H. tuberculata Hth
17	38	66.7	414	21 AAB10995	H. tuberculata hem
18	38	66.7	414	22 AAB71107	H. tuberculata Hth
19	38	66.7	415	21 AAB10977	H. tuberculata hem
20	38	66.7	415	22 AAB71089	H. tuberculata Hth
21	37	64.9	415	22 AAB71145	H. tuberculata Hth
22	37	64.9	422	15 AAB63186	Variant alpha amy1
23	37	64.9	468	13 AAR24136	Alpha-amyase vari
24	37	64.9	468	15 AAR63187	Variant alpha amy1
25	37	64.9	468	15 AAR63184	Variant alpha amy1
26	37	64.9	468	15 AAR63185	Variant alpha amy1
27	37	64.9	494	8 AAP70571	Alpha-amyase gene
28	37	64.9	4820	22 ABB58592	Drosophila melanog
29	36	63.2	42	22 AAB68777	Human immune/hema
30	36	63.2	165	21 AAG33468	Arabidopsis thalia
31	36	63.2	198	21 AAG33467	Arabidopsis thalia
32	36	63.2	266	21 AAG33466	Arabidopsis thalia
33	36	63.2	514	21 AAB42398	Human ORFX ORF2162
34	36	63.2	717	22 AAT19456	Human diagnostic a
35	35	61.4	84	22 AAB03143	Human gene 21 enco
36	35	61.4	121	22 AAB02815	Protonibacterium
37	35	61.4	162	22 AAB03117	Human gene 21 enco
38	35	61.4	255	20 AAG31833	Human foetal kidn
39	35	61.4	255	22 ABB02357	Novel human diagno
40	35	61.4	255	22 AAB29292	Human PRO polypept
41	35	61.4	255	22 AAB73684	Human oxidoreducta
42	35	61.4	349	20 AAT09298	Rice beta-glucanasa
43	35	61.4	377	22 AAG07906	Rice Gns9 CDS prot
44	35	61.4	390	22 AAB67596	Transcription fact
45	35	61.4	415	21 AAG21556	Arabidopsis thalia

ALIGNMENTS

RESULT 1
ID AAR46065 standard; protein; 478 AA.
AC AAR46065;
DT 18-JUL-1994 (first entry)
XX
DE Mutant alpha-amyase.
XX
KW Methionine substitution; stability; activity; detergent;
KW dishwashing agents; liquidation agents.
XX
XX Aspergillus oryzae.
OS
XX
XX
XX W09402597-A.
XX
XX 03-FEB-1994.
XX
XX
XX 06-JUL-1993; 93WO-DK00230.
XX
XX 23-JUL-1992; 92DK-0000946.
XX 16-DEC-1992; 92DK-0001503.
XX 15-MAR-1993; 93DK-0000292.
XX
XX (NOVO) NOVO-NORDISK AS.
XX
XX Bisgard-frantzen H, Svendsen A;
XX WPI. 1994-048855/06.
XX
XX Mutant alpha-amyase from Bacillus species comprising a
XX methionine substitution - with improved stability and activity at
XX low pH, for use in detergents, dishwashing agents and

PT Liquefaction agents
XX
PS Claim 1; Page 7; 20pp; English.
XX
CC The sequence of that of the *Aspergillus oryzae* alpha amylase, sold
CC commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can
CC be mutated by substitution of one or more of its methionine residues
CC for any amino acid other than cysteine. The mutant alpha-amylase
CC exhibits a better activity level and better stability in the
CC presence of oxidising agents than previous mutant alpha amylases,
CC and improved thermostability at moderately low pH. The enzyme can
CC be used as an additive for detergents, dishwashing agents and
CC liquidation agents.
XX
SQ Sequence 478 AA:

Query Match 100.0%; Score 57; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFEKPF 11
130 ssvdysvfkpf 140

RESULT 2

AAR79025 standard; protein; 478 AA.

AAR79025;

22-MAR-1996 (first entry)

Mature taka-amylase A.

Wild type; neopullulanase; B. stearotherophilus; mutant;
food industry; modification; hydrophobicity; replacement
insertion; deletion.

Aspergillus oryzae.

Location/Qualifiers

Disulfide-bond 30..38

Disulfide-bond 150..164

Disulfide-bond 240..283

Disulfide-bond 439..474

JP07177891-A.

18-JUL-1995.

31-OCT-1994; 94JP-0288658.

12-NOV-1993; 93JP-0306096.

(EZAKI) EZAKI GLICO CO.

(NIDE) NEC CORP.

WPI; 1995-279919/37.

Modifying a transferase by enhancing hydrophobicity of a selected
site - increases transfer activity, also new mutant
neo-pullulanase(s)

Disclosure: Page 10-11; 18pp; English.

This sequence represents the mature form of taka-amylase from *A. oryzae*.
This sequence was used in a method for the generation of mutant
pullulanases for use in the food industry (see also AAR79026-28). The
wild type pullulanase enzyme was modified by the method of the
invention for enhancing the hydrophobicity of a selected site of the
pullulanase. The method comprises replacement of a group in the

selected site with a hydrophobic group, replacement of an amino acid
with a hydrophobic amino acid, and/or insertion or deletion of a
hydrophobic amino acid from the selected site. The method was used
to produce neopullulanases Y377F, S422V and M375L.

Sequence 478 AA:

Query Match 100.0%; Score 57; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFEKPF 11
130 ssvdysvfkpf 140

RESULT 3

AAR72450 standard; Protein; 478 AA.

AAR72450;

01-DEC-1995 (first entry)

Aspergillus oryzae alpha amylase (mature protein).

Alpha amylase; variant; enzyme; detergent; additive; dishwashing;

washing; Bacillus licheniformis; Bacillus amyloliquefaciens;

Bacillus stearotherophilus; dyeing; bleaching; scouring; textile;

thermostable.

Aspergillus oryzae.

WO9510603-A.

20-APR-1995.

05-OCT-1994; 94WO-DK00370.

08-OCT-1993; 93DK-0001133.

02-FEB-1994; 94DK-0000140.

(NOVO) NOVO-NORDISK AS.

Svendsen A, Thellersen M;

Van der zee P, Bisgaardfrantzen H, Borchert T;

WPI; 1995-161790/21.

New Bacillus derived alpha-amylase variants - having amino acid

modifications to improve washing and/or dishwashing performance

Disclosure: Page 75-76; 105pp; English.

Variant alpha amylase enzymes which have improved washing and/or

as detergent additives. The enzymes have one or more amino acid

residues added, deleted or substituted. The variants can also be

used for textile desizing prior to scouring, bleaching and dyeing.

The variants have improved thermostability, acid/alkaline stability;

low temperature optimum; pH optimum; higher hydrolysis velocity and

improved tolerance to other composition constituents, e.g. oxidation

agents.

Sequence 478 AA:

Query Match 100.0%; Score 57; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVDSVFEKPF 11
130 ssvdysvfkpf 140

Db 130 ssvdysvfkpf 140

RESULT 4
AAR78270
ID AAR78270 standard; Protein; 478 AA.
XX
XX AAR78270;
AC
XX
XX 17-JAN-1996 (first entry)
DE Aspergillus oryzae alpha amylase (mature protein).
XX
XX Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;
KM starch; thermostable; methionine; Bacillus licheniformis;
KM Bacillus thermoliquefaciens; Bacillus steothermophilus;
XX Aspergillus oryzae.
XX
XX Aspergillus oryzae.
XX
XX W09521247-A1.
XX
XX 10-AUG-1995.
XX
XX 05-OCT-1994; 94MO-DK00371.
XX
XX 02-FEB-1994; 94DK-0000141.
XX
XX (NOVO) NOVO-NORDISK AS.
XX
XX Marcher D, Nilsson TE, Pedersen HH, Toft AH;
XX
XX WPI, 1995-283767/37.
XX
XX Use of an oxidation stable alpha-amylase - for simultaneous desizing
PT and bleaching or scouring of fabrics contg. starch or starch derivs.
XX
XX Disclosure; Page 25-26; 37pp; English.
XX
XX
XX Oxidation stable alpha amylases can be used for the simultaneous
CC desizing and bleaching or scouring of a fabric comprising starch or
CC starch derivatives. They exhibit a better heat stability, and
CC especially in the presence of oxidizing agents. They are obtained
CC from a parent alpha amylase by replacing one or more methionine
CC residues with any amino acid different from Cys or Met, preferably
CC Leu, Thr, Ala, Gly, Ser, Ile or Asp. The parent alpha amylase is
CC pref. derived from a Bacillus species, although alpha amylases of
CC fungal origin can also be used. This sequence is the wild type
CC (unmodified) alpha amylase of Aspergillus oryzae.
XX
XX
SQ Sequence 478 AA;

Query Match 100.0%; Score 57; DB 16; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSVYDYSVFKPF 11
XXXXXXXXXXXX
Db 130 ssvdysvfkpf 140

RESULT 5
AAW14500
ID AAW14500 standard; protein; 478 AA.
XX
XX AAW14500;
AC
XX
XX 04-JUN-1997 (first entry)
DE Aspergillus oryzae alpha-amylase (mature protein).
XX
XX alpha-amylase; mature; Bacillus licheniformis; termamyl; fungamyl;

KM Aspergillus oryzae; Bacillus amyloliquefaciens; altered property;
KM calcium dependency; substrate binding; stability; pH optimum;
KM thermostability; cleavage; oligosaccharide substrate; dishwashing;
KM washing; detergent additive; fabric desizing; starch liquefaction;
KM sweetener; ethanol production; variant.
XX
XX Aspergillus oryzae.
OS
FH Location/Qualifiers
FH
FH Key
FH
FH Misc-difference 13..45
FT /label= loop 1 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 7-23 of AAW14499 is deleted or
FT replaced with a fragment corresponding to
FT this fragment; claim 33"
FT
FT
FT Misc-difference 14..40
FT /label= loop 1 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 8-18
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 35"
FT
FT Misc-difference 28..42
FT /label= loop 1 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 12-19 of AAW14499 is deleted or
FT replaced with a fragment corresponding to
FT this fragment; claim 30"
FT
FT
FT Misc-difference 32..38
FT /label= loop 1 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 14-15
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 32"
FT
FT Misc-difference 66..84
FT /label= loop 2 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 44-57 of AAW14499 is deleted or
FT replaced with a fragment corresponding to
FT this fragment; claim 18"
FT
FT
FT Misc-difference 70..78
FT /label= loop 2 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 48-51
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 20"
FT
FT Misc-difference 98..210
FT /label= loop 3 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 117-185 of AAW14499 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 24"
FT
FT
FT Misc-difference 102..206
FT /label= loop 3 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 121-181
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 26"
FT
FT Misc-difference 121..174
FT /note= "preferred region where at least one amino acid

FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to this
FT fragment is deleted or replaced with a fragment
FT corresponding to 102-199 of AAM14498; claim 42"

FT Misc-difference 165..177
FT /label="loop 3 modification region
FT /note="at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 195-202 of AAM14499 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 21"

FT Misc-difference 166..173
FT /label="loop 3 modification region
FT /note="preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 196-198
FT of AAM14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 23" Misc-difference 181..184
FT /note="an amino acid fragment corresponding to this
FT region is deleted from the parent sequence of a
FT variant Fungamyl; claim 43"

FT Misc-difference 291..313
FT /label="loop 8 modification region
FT /note="at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 322-346 of AAM14498 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 36"

FT Misc-difference 297..313
FT /label="loop 8 modification region
FT /note="preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 325-345
FT of AAM14498 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 38"

PN MO9623874-A1.
XX
XX 08-AUG-1996.
PD
XX 05-FEB-1996; 96WO-DK00057.
PR
XX 10-NOV-1995; 95DK-0001236.
PR 03-FEB-1995; 95DK-0000128.
PR 23-OCT-1995; 95DK-0001192.
XX
XX (NOVO) NOVO-NORDISK AS.
XX
XX Bisgard-frantzen H, Borchert TV, Svendsen A;
XX WPI; 1996-371424/37.
XX
XX Alpha-amylase variants and methods of production - have altered
XX properties such as calcium dependency, substrate binding and
XX stability
XX
XX Disclosure; Page 87-88; 171pp; English.
XX
XX The present sequence is the mature Aspergillus oryzae alpha-amylase (A.).
XX variants of parent termamyl-1 and fungamyl-1-like alpha-amylases (and
XX methods of constructing them) are claimed. Examples of variants are
XX featured above. The variants have altered properties such as calcium
XX dependency, substrate binding and stability. Also one or more proline or
XX cysteine residues in the variant is modified or replaced with a
XX non-proline or non-cysteine residue such as alanine. The variants can be
XX used for (dish)washing, as detergent additives or for fabric desizing or
XX starch liquefaction. They can also be used for the production of
XX sweeteners and ethanol from starch. See also AAM14498-99.
XX
XX Sequence 478 AA;

```

QY      1  SVDYSVFKPF 11      100.0%: Score 57; DB 17; Length 478;
        |||
DB      130  ssvdysvfkpf 140      100.0%: Pred. No. 0.0062;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0

RESULT 6
AAB84206
ID      AAB84206 standard; Protein; 498 AA.
AC      AAB84206;
XX
XX      06-AUG-2001 (first entry)
DE
XX      Amino acid sequence of a fungamyl-like alpha-amylase.
KW      Fungamyl-like alpha-amylase; glucamylase; dextrinisation; maltose;
KN      alcohol; starch; dough improver; brewing; starch liquification.
XX
XX      Aspergillus oryzae.
OS
XX      WO200134784-A1.
PN
PD      17-MAY-2001.
PE      10-NOV-2000; 2000WO-DK00626.
XX
XX      10-NOV-1999; 99DK-0001617.
FR
XX      (NOVO ) NOV02YMES AS.
PA
XX
XX      Bisgard-Frantzen H, Svendsen A, Pedersen S;
PI
XX      WPI: 2001-367478/38.
DR
XX      N-PSDB; AAF90208.
PT
XX
XX      New variant of Fungamyl-like alpha-amylase, useful for production of
PT      maltose syrups, includes mutations that improve stability against heat
PT      and acidic pH
PS
XX      Claim 1; Page 42-45; 49pp; English.
XX
XX
XX      The present sequence represents a fungamyl-like alpha-amylase. The
CC      specification describes variants of this fungamyl-like alpha-amylase,
CC      which have an alteration in one the amino acid regions 98-110, 150-160,
CC      161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion
CC      or substitution of an amino acid or an insertion of an amino acid
CC      downstream of a particular position. The variants retain alpha-amylase
CC      activity, and have better heat stability and/or stability at acidic pH,
CC      relative to wild-type enzyme. The variants can therefore be used at
CC      higher temperatures (more efficient conversion or faster reaction), and
CC      have reduced need for cooling and reduced risk of contamination). The
CC      variants may also be used in conjunction with other enzymes,
CC      particularly glucoamylase during dextrinisation. The variants are
CC      used to produce syrups, particularly of high maltose content, or alcohol,
CC      from starch, as dough improver for baked goods; in brewing, to increase
CC      fermentability of the wort; and for liquifaction of starch.
XX
XX      Sequence 498 AA;
SO

Query Match 100.0%; Score 57; DB 22; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  SVDYSVFKPF 11
        |||
DB      150  ssvdysvfkpf 160

```

```

RESULT      7
AAP81161
ID   AAP81161 standard; protein; 511 AA.
AC     AAP81161;
XX
DT    12-OCT-1990 (first entry)
XX
DE    Recombinant alpha-amylase.
XX
KW    Alpha amylase; yeast; brewing; baking.
XX
OS    Schwanniomycetes castellii ATCC 26076.
XX
PN    EP257115-A.
XX
02-MAR-1988.
XX
FF    21-AUG-1986; 86EP-0111586.
XX
PR    21-AUG-1986; 86EP-0111586.
XX
PA    (HETB ) HEINEKEN TECHNISCH BEHEER BV.
XX
PI    Strasser A, Martens FB, Dolmen J, Hollenberg CP;
DR    WPI; 1988-057532/09.
DR    N-PSDB; AAN81477.
XX
PT    Prodn. of amylolytic enzymes by transformed microorganisms -
XX    comprises use of recombinant DNA technology to give transformant
XX    having DNA from donor yeast.
XX
PS    Disclosure; ; P: English.
XX
The sequence encoding the protein can be used to construct an
expression vector for the prodn. of alpha amylase in yeast cells.
The transformants and recombinant enzyme produced can be used in
fermentation processes e.g in baking and brewing. The transformed
yeasts are esp. suitable for prodn. of low carbohydrate beers.
XX
SQ    Sequence      511 AA;

Query Match          78.9%; Score 45; DB 9; Length 511;
Best Local Similarity 81.8%; Pred. No. 1.6;
Matches      9; Conservative      0; Mismatches      2; Indels      0; Gaps      0;

QY      1 SSVDSVSKPPE 11
        ||||| | ||
Db       165 ssvdysstfpf 175

RESULT      8
AAP81180
ID   AAP81180 standard; protein; 511 AA.
AC     AAP81180;
XX
DT    12-DEC-1990 (first entry)
XX
DE    Sequence of alpha-amylase encoded by genomic DNA of Schwanniomycetes
DE    castellii.
XX
KW    Brewing; beer; breadmaking; biomass.
XX
OS    Schwanniomycetes castellii.
XX
PN    EP260404-A.
XX
PD    23-MAR-1988.
XX

```

PF	17-JUL-1987;	87EP-0110370.
XX		
PR	17-JUL-1987;	87EP-0110370.
XX		
PA	(HETB) HEINEKEN TEC BEHEER NV.	
XX		
PI	Strasser A, Martens FB, Dohmen J, Hollenberg CP;	
XX		
DR	WPI: 1988-078794/12.	
XX		
PT	esp. coding alpha amylase or glucosylase, able to convert	
PT	starch to ethanol	
XX		
PS	Claim 18, Fig 2A-2C; 74pp; English.	
XX		
CC	A cosmid library is constructed from genomic DNA of ATCC26076 and this	
CC	screened for inserts contg. the AA gene by ability to transform	
CC	S.cerevisiae GPF 18 to histidine prototrophy. One positive cosmid	
CC	(3kb). pYCl1 was digested with EcoRI, religated and used to transform	
CC	E.coli DH221. Positive transformants contain pYCl1-alpha, which includes	
CC	a 5kb EcoRI fragment contg. the gene in AAN61525.	
XX		
SO	Sequence 511 AA;	
	Query Match 78.9%; Score 45; DB 9; Length 511;	
	Best Local Similarity 81.8%; Pred. No. 1.6;	
	Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0.	
OY	1 SSVDSYVKEPF 11	
DB	165 ssvdgsstpf 175	
	RESULT 9	
	AAR07574	
ID	AAR07574 standard; protein: 512 AA.	
XX		
AC	AAR07574;	
XX		
DT	02-FEB-1991 (first entry)	
XX		
DE	Alpha-amylase encoded by the AMY1 gene.	
XX		
KW	Schwanniomyces yeast cells; expression cassette; alpha-amylase;	
KW	marker gene; regulation: signal peptide; terminator;	
KM	autonomously replicating sequence.	
XX		
OS	Schwanniomyces occidentalis.	
XX		
FH	Key Location/Qualifiers	
FT	Disulfide-bond 66..74	
FT	/note="1"	
FT	Disulfide-bond 186..200	
FT	/note="2"	
FT	Disulfide-bond 276..319	
FT	/note="3"	
FT	Disulfide-bond 475..510	
FT	/note="4"	
XX		
PN	EP394538-A.	
XX		
PD	31-OCT-1990.	
XX		
PF	28-APR-1989; 89EP-0107780.	
XX		
PR	28-APR-1989; 89EP-0107780.	
XX		
PA	(RHEI) RHEIN BIOTECH GES.	
XX		
PI	Hollenberg C, Strasser A;	
XX		

```

XX WPI; 1990-328670/44.
DR N-PSDB; AA006388.
XX
PT Transformed Schwanniomycetes yeast cells - contg. an expression
PT cassette confg. regulon, DNA coding for foreign protein and
XX terminator
XX
PS Disclosure: Fig 1B(1-6); 59pp; English.
XX
CC The expression cassette may contain a regulon (R), a signal peptide
CC sequence (S), a foreign protein sequence (P) and a terminator (T)
CC of which R and/or S and/or T are derived from the alpha-amylase gene
CC of S. occidentalis.
CC R is pref. a 1.8 kb BglII-XhoI fragment (bases -1 to -540 of this
CC sequence).
CC S contains all or part of one of the sequences represented
CC in AA006383-87.
CC F is e.g. a cellulase, interleukin, insulin-like-growth factor,
CC interferon etc.
CC T is pref. all or part of the terminator of this sequence
CC (bp 1537-1740).
CC AMY1 acts as a selective marker for the vector carrying the expression
CC cassette.
CC See also AA006389.
XX
SQ Sequence 512 AA:

Query Match 78.9%; Score 45; DB 11; Length 512;
Best Local Similarity 81.8%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSVDYSVFKPF 11
Db 166 ssvdysstfcpf 176
|||||
|

RESULT 10
AAR88212
ID AAR88212 standard; Protein; 493 AA.
XX
AC AAR88212;
XX
DT 03-APR-1996 (first entry)
XX
DE Alpha-amylase.
XX
KW Alpha-amylase; thermostable enzyme; baking; Aspergillus niger.
XX
Thermomyces lanuginosus CBS 224.63.
XX
FT Key Location/Qualifiers
FT Peptide 1..18
FT /label= Sig_peptide

WO9601323-A1.
XX
PN 18-JAN-1996.
XX
PD 03-JUL-1995; 95WO-EP02607.
XX
PF 04-JUL-1994; 94GB-0013419.
XX
PR (DANI-) DANISCO AS.
XX
PA Michelsen B, Rasmussen P;
XX WPI; 1996-087673/09.
XX DR N-PSDB; AAT10562.
XX
PT Thermophilic alpha-amylase with activity range of 60-80 degrees C -
PT derived from Thermomyces lanuginosus, useful in the prepn. of

```

```

PT foodstuffs and bakery prods. esp. bread
XX
PS Claim 3; Page 36-38; 94pp; English.
XX
CC A thermostable alpha-amylase (AAR88212) of Thermomyces lanuginosus
CC CBS 224.63 is expressed in Aspergillus niger hosts using a gene
CC (AAT10562) isolated from a T. lanuginosus gene library. The
CC recombinant enzyme (54-60 kDa) shows optimal activity at
CC 60-70 deg and pH 5.8-6, has a pI of 3.7 and is active at 60-80
CC deg.
XX
SQ Sequence 493 AA:

Query Match 71.9%; Score 41; DB 17; Length 493;
Best Local Similarity 70.0%; Pred. No. 9.4;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSVDYSVFKPF 10
Db 150 ssvdysstfcp 159
|||||
|

RESULT 11
AAV53678
ID AAV53678 standard; Protein; 5250 AA.
XX
AC AAV53678;
XX
DT 22-FEB-2000 (first entry)
XX
DE Sequence g1/4426611/gb/AAD204501 from an alignment with protein 274.
XX
KW Mechanical stress; gene therapy; protein 274; osteoporosis; bone density;
KW bone development; g1/3328186.
XX
OS unidentified.
XX
PN WO9960164-A1.
XX
PD 25-NOV-1999.
XX
PF 14-MAY-1999; 99WO-US11066.
XX
PR 15-MAY-1998; 98US-0085673.
XX
PA (QUAR-) QUARK BIOTECH INC.
XX
PI Einat P, Mor O, Skallier R, Feinstein E, Faerman A;
XX WPI; 2000-053304/04.
XX
PT Identification of stress induced genes for determining risk and
PT preventing, treating or controlling osteoporosis
XX
PS Claim 32; Fig 14A-T; 308pp; English.
XX
CC The present sequence is obtained from a clustal X alignment with
CC protein 274. Protein 274 was identified using the method of the invention
CC after subjecting rat osteoblasts to mechanical stress. Expression of the
CC 608 gene was found to be upregulated by about 3-fold in cells subjected
CC to mechanical strain. The specification describes a method for the
CC identification of genes responsive to a specific mechanical stress. The
CC method comprises applying the mechanical stress to an organism (tissue
CC or cells comprising bone cells), isolating the specific cellular
CC fractions and extracting mRNA from them, and differentially analyzing the
CC mRNA in comparison with control samples. The method is used to identify
CC genes whose expression is responsive to a specific stress. The identified
CC genes are employed in determining risk associated with a physiological or
CC disease state. The risk determination methods are used for testing a
CC medicament for gene therapy. These medicaments, or genes identified by
CC the method of the invention, are used for treating, preventing or
CC controlling a physiological or disease state (especially osteoporosis or

```

CC bone density or other factors causing or contributing to osteoporosis or
CC its symptoms or other conditions involved in mechanical stress or its
CC lack. The methods can also be used for advancing research or studies in
CC bone development.

XX Sequence 5250 AA;

Query Match 68.4%; Score 39; DB 21; Length 5250;

Best Local Similarity 75.0%; Pred. No. 3.7e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 DYSVEKPF 11

Db 5147 dytykpf 5154

WU 12

ID ABB69422 standard; Protein; 5322 AA.

XX ABB69422;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 35058.

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

WPI: 2001-656860/75.

N-PSDB; ABL13525.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

PS Disclosure; SEQ ID NO 35058; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AAB57737-AB572072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 5322 AA;

Query Match 68.4%; Score 39; DB 22; Length 5322;

Best Local Similarity 75.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 DYSVEKPF 11

Db 5219 dytykpf 5226

RESULT 13

AAV41363

ID AAV41363 standard; Protein; 55 AA.

AC AAV41363;

DT 02-DEC-1999 (first entry)

DE Human secreted protein encoded by gene 56 clone HNGFE55.

KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

PN WO9947540-A1.

PD 23-SEP-1999.

PF 18-MAR-1999; 99WO-US05804.

PR 19-MAR-1998; 98US-0078563.

PR 19-MAR-1998; 98US-0078566.

PR 19-MAR-1998; 98US-0078573.

PR 19-MAR-1998; 98US-0078574.

PR 19-MAR-1998; 98US-0078576.

PR 19-MAR-1998; 98US-0078577.

PR 19-MAR-1998; 98US-0078578.

PR 19-MAR-1998; 98US-0078579.

PR 19-MAR-1998; 98US-0078581.

PR 01-APR-1998; 98US-0080312.

PR 01-APR-1998; 98US-0080313.

PR 01-APR-1998; 98US-0080314.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;

PI Wei Y, Endress GA, Duan RD, Kyaw H, Eoner R, Lafleur DW,

PI Olsen HS, Shi Y, Moore PA;

WPI: 1999-562050/47.

N-PSDB; AA224866.

PT New isolated human genes, useful for diagnosis and treatment of e.g.
PT cancers, neurological disorders, immune diseases, inflammation or blood
PT disorders -

PS Claim 11; Page 391; 484pp; English.

XX This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AA224802) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 95 novel genes and their fragments (nucleic
CC acid sequences: AA224811-224907; amino acid sequences AAV41308-Y41404)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 95
CC polynucleotides, based on which tissues they are most highly expressed in

CC (see AAZ24811 for described uses).

XX Sequence 55 AA;

SO

Query Match
Best Local Similarity 66.7%; Score 38; DB 20; Length 55;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SSVDSYVFKP 10
|||:|||||
Db 21 ssvfsvfkp 30

RESULT 14

AAB10970
ID AAB10970 standard; Protein: 413 AA.

AC AAB10970;

DT 16-FEB-2001 (first entry)

XX H. tuberculata hemocyanin Hth1 domain d.

KM Hemocyanin; cytosolic; virucide; antibacterial; antiparasitic;
KM immunomodulatory; antihypertensive; gene therapy; tumor; treatment;
KM infection; schistosomiasis; carcinoma; bladder; epithelium; ovary;
KM breast; bronchi; colon-rectum; hypertension; vaccine; cocaine misuse;
KM pharmaceutical carrier.

XX Haliotis tuberculata.

OS WO20005192-A2.

XX 21-SEP-2000.

XX 17-MAR-2000; 2000WO-EP02410.

XX 17-MAR-1999; 99DE-1011971.

XX 20-AUG-1999; 99DE-1039578.

XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.

XX Markl J, Altenhein B, Lieb B, Stiefel T;

XX WPI: 2000-587517/55.

XX New nucleic acid encoding hemocyanin, useful for gene therapy of tumors
XX and for recombinant production of fusion proteins for vaccination -

XX Claim 21; Page 91-93; 163pp; German.

CC This invention describes a novel nucleic acid (I) containing a sequence
CC that encodes hemocyanin (II), a domain of (I) or its fragment with the
CC immunological properties of at least one domain of (II). The products of
CC the invention have cytosolic, virucide, antibacterial, antiparasitic,
CC immunomodulatory and antihypertensive activity. (II), and constructs
CC additionally containing antigen-encoding sequences, are useful in gene
CC therapy of tumors. Polypeptides encoded by (I) are useful for treating
CC parasitic or viral infections and tumors, particularly schistosomiasis
CC and carcinoma (of bladder, epithelium, ovary, breast, bronchi or
CC colon-rectum), also hypertension, as vaccines, for treating cocaine
CC misuse and very generally as carriers for pharmaceuticals, e.g.
CC cytostatics. They may also be used to generate antibodies (Ab). Probes
CC based on (I) and Ab are useful for detecting tumor-specific DNA in a cell
CC (by detecting specific binding to cellular DNA or proteins), particularly
CC where associated with the types of carcinoma listed above. Hemocyanins
CC can be produced recombinantly, relatively inexpensively and in adequate
CC amounts, eliminating the need to culture gastropods. When used as a
CC carrier, (II) significantly increases the half-life of the attached
CC pharmaceutical, by inhibiting ultrafiltration in the kidneys.

XX Sequence 413 AA;

SO

Query Match
Best Local Similarity 66.7%; Score 38; DB 21; Length 413;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSVDSYVFKP 10
|||:|||||
Db 200 sldysatdp 209

RESULT 15

AAB71082
ID AAB71082 standard; Protein: 413 AA.

AC AAB71082;

DT 24-AUG-2001 (first entry)

XX H. tuberculata Hth1 domain d.

KM Hth1; KLM1; hemocyanin; cytosolic; vasotropic; vaccine; gene therapy;
KM tumor; parasite infection; viral infection; antitumor agent;
KM abnormal blood pressure; bladder carcinoma; epithelial carcinoma;
KM ovarian carcinoma; mammary carcinoma; bronchial carcinoma;
KM colon carcinoma; cocaine addiction.

OS Haliotis tuberculata.

XX WO200114536-A2.

XX 01-MAR-2001.

XX 21-AUG-2000; 2000WO-EP08129.

XX 20-AUG-1999; 99DE-1039578.

XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.

XX Markl J, Altenhein B, Lieb B, Stiefel T;

XX WPI: 2001-191646/19.

XX New nucleic acid encoding hemocyanin, useful for gene therapy of tumors
XX and for recombinant production of fusion proteins for vaccination -

XX Claim 25; Page 108-110; 180pp; German.

CC This invention describes a novel nucleic acid molecule (N1) encoding
CC hemocyanin, a hemocyanin domain, or a functional fragment of it with
CC the immunological properties of at least one domain of hemocyanin, and
CC which comprises at least one intron is new. The products of the invention
CC have cytostatic and vasotropic activity and can be used in vaccines or
CC for gene therapy. Compositions comprising nucleic acids of the invention
CC are useful in treating tumors. Compositions comprising hemocyanin
CC polypeptide are useful in treating parasite and viral infections and as
CC an antitumor agent. It is also useful in treating abnormal blood
CC pressure, bladder carcinomas, epithelial carcinomas, ovarian carcinomas,
CC mammary carcinomas, bronchial carcinomas, and colon carcinomas. It may
CC also be used to cure cocaine addiction.

XX Sequence 413 AA;

SO

Query Match
Best Local Similarity 66.7%; Score 38; DB 22; Length 413;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSVDSYVFKP 10
|||:|||||
Db 200 sldysatdp 209

Fri Jun 14 10:36:41 2002

us-09-710-339-2_copy_150_160.rag

Page 9

Search completed: June 13, 2002, 08:53:26
Job time: 276 sec

•
•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:49:40 ; Search time 73.31 Seconds
(without alignments)
652.742 Million cell updates/sec

Title: US-09-710-339-2
Perfect score: 2684
Sequence: 1 MVWMSLFLYGLQVAPALA.....LPVLYPTKLAGSKICSS 498

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Indexed: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	2684	100.0	499	1	ALAS1
2	2675	99.7	499	2	JS0663
3	2674	99.6	499	1	ALAS3
4	2674	99.6	499	2	B48305
5	2673	99.6	499	2	JN0588
6	2662	99.2	488	2	A48305
7	2656	99.0	499	2	JT0466
8	2503.5	93.3	478	2	JK0201
9	1780	66.3	484	1	A35282
10	1629	60.7	624	1	JC4510
11	1527	56.9	507	2	S33921
12	1449	54.0	512	2	S06115
13	1445.5	53.9	512	2	S23355
14	1397.5	52.1	484	1	ALB1AF
15	1264	47.1	631	1	S72270
16	1015.5	37.8	513	2	T38770
17	963.5	35.9	625	2	T41603
18	920.5	34.3	491	2	T38448
19	899	33.5	581	2	S62505
20	898.5	33.5	564	2	T41503
21	898	33.5	478	2	T40860
22	737	27.5	482	2	S31478
23	670	25.0	1196	2	A29130
24	653.5	24.3	774	2	T39539
25	481.5	17.9	642	2	A11827
26	456.5	17.0	713	1	ALBSG1
27	454.5	16.9	712	1	ALBSG3
28	444.5	16.6	713	1	ALBSG7
29	440.5	16.4	528	1	ALBSK

30	433	16.1	710	2	S63598	cyclomaltoedextrin
31	428.5	16.0	655	1	ALBSG	cyclomaltoedextrin
32	428	15.9	703	1	ALBSX1	cyclomaltoedextrin
33	424.5	15.8	919	2	S28179	alpha-amylose (EC
34	418	15.6	713	2	A58800	cyclomaltoedextrin
35	418	15.6	714	1	ALBSG6	cyclomaltoedextrin
36	414.5	15.4	713	2	S09196	cyclomaltoedextrin
37	412	15.4	717	1	S28784	alpha-amylose (EC
38	410	15.3	483	2	G75392	glycosyl hydrolase
39	407	15.2	704	2	T39805	cyclomaltoedextrin
40	406.5	15.1	718	1	ALBSG6	cyclomaltoedextrin
41	405.5	15.1	718	1	ALBSG6	cyclomaltoedextrin
42	405	15.1	713	1	ALBSMX	cyclomaltoedextrin
43	399.5	14.9	718	1	ALBSMX	cyclomaltoedextrin
44	384.5	14.3	711	1	ALBSXF	cyclomaltoedextrin
45	376.5	14.0	1104	2	A60999	alpha-amylose (EC

ALIGNMENTS

RESULT 1
ALAS1
alpha-amylose (EC 3.2.1.1) 1 precursor [validated] - Aspergillus oryzae
N:Alternate names: alpha-amylose isozyme II; glycogenase; Taka-amylose A
C:Species: Aspergillus oryzae
C>Date: 30-Sep-1991 #sequence-revision 30-Sep-1991 #text-change 15-Sep-2000
C:Accession: S04548; A33214; JS0240; A91930; A93767; A10627
R:Wise, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
Mol. Microbiol. 3, 3-14, 1989
A>Title: Three alpha-amylose genes of Aspergillus oryzae exhibit identical intron-exo
A:Reference number: S04548; MUID:89237897
A:Accession: S04548
A:Molecule type: DNA
A:Residues: 1-499 <MIR>
A:Cross-references: EMBL:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921
A:Genetics: AMY1
A:Accession: A33214
A:Molecule type: mRNA
A:Residues: 1-499 <MIR>
A:Cross-references: GB:X12725; NID:g2430; PIDN:CAA31218.1; PID:g295921
R:Genes, M.J.; Dove, M.J.; Seligy, V.L.
Gene 79, 107-117, 1989
A>Title: Aspergillus oryzae has two nearly identical Taka-amylose genes, each contain
A:Reference number: JS0240; MUID:89378767
A:Accession: JS0240
A:Molecule type: DNA
A:Residues: 1-499 <GEN>
A:Genetics: AMY2
A>Note: the authors refer to this as isozyme II
R:Isemura, S.; Ikenaka, T.
J. Biochem. 74, 1-10, 1973
A:Reference number: A91930; MUID:74001521
A:Accession: A91930
A:Molecule type: cDNA
A:Residues: 206-225 <ISE>
R:Narita, K.
Proc. Jpn. Acad. 51, 285-290, 1975
A:Reference number: A93767
A:Accession: A93767
A:Molecule type: protein
A:Residues: 434-443,446-447, 'O',449-458, 'GTVV',459-464,467-468, 'B',470, 'B',472-499 <N
R:Watanabe, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
J. Biochem. 95, 697-702, 1984
A>Title: Structure and possible catalytic residues of Taka-amylose A.
A:Reference number: A37454; MUID:84212370
A:Contents: annotation; X-ray crystallography, 3.0 angstroms
R:Swift, H.J.; Brady, L.; Derewenda, Z.S.; Dodson, E.J.; Turkunburg, J.P.; Wilkinson,
submitted to the Brookhaven Protein Data Bank, August 1992
A:Reference number: A51548; PDB:6TAA
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 22-497
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics: <AMY1>

A:Gene: amy1
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Genetics: <AMY2>
 A:Gene: amyII
 A:Introns: 56/3; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycoen/starch degradation
 C:Superfamily: Aspergillus alpha-amyase; alpha-amyase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <Sig>
 F:22-499/Product: alpha-amyase #status predicted <AMP>
 F:194-321/Domain: alpha-amyase core homology <AMY>
 F:51-59,171-185,261-304,461-496/Disulfide bonds: #status experimental
 F:142,183,196,231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:227,251,318/Active site: Asp, Glu, Asp #status predicted

231 21 323

Query Match 100.0%; Score 2684; DB 1: Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.2e-185;
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1  MVAWMSLFLVGLQVAAAPALATPADMRSSOSIFLLTDFARTDSTTATCTADOKXCG 60
   |||||||
2  MVAWMSLFLVGLQVAAAPALATPADMRSSOSIFLLTDFARTDSTTATCTADOKXCG 61
   |||||||
61  TWGGIIDKLDYIOGMGFTAIWTPVTAQLPQTAYGDAYHGWOODIYSLNENGTADL 120
   |||||||
62  TWGGIIDKLDYIOGMGFTAIWTPVTAQLPQTAYGDAYHGWOODIYSLNENGTADL 121
   |||||||
121  KALSSALHERGMVLMVDVYANMHMGYDAGSSVDYSVFRPSSQDYFHPFCFIONYEOTO 180
   |||||||
122  KALSSALHERGMVLMVDVYANMHMGYDAGSSVDYSVFRPSSQDYFHPFCFIONYEOTO 181
   |||||||
181  VEDCWIGDNTVSLPDLDTTKDYVKNEMWDVGSVSNYSIDGLRIDYKHKQDFWPGYN 240
   |||||||
182  VEDCWIGDNTVSLPDLDTTKDYVKNEMWDVGSVSNYSIDGLRIDYKHKQDFWPGYN 241
   |||||||
241  KAAGVYCIGEVLDGDPATYCPYQNVMDGVLPYIYPLLNKFKSTSGSMDLLYNNINTVK 300
   |||||||
242  KAAGVYCIGEVLDGDPATYCPYQNVMDGVLPYIYPLLNKFKSTSGSMDLLYNNINTVK 301
   |||||||
301  SDCPSTLLGTFFVENHNDREFASYTNDIALAKNVAAFIILNDGIPPIYAGOEORHAGND 360
   |||||||
302  SDCPSTLLGTFFVENHNDREFASTINDIALAKNVAAFIILNDGIPPIYAGOEORHAGND 361
   |||||||
361  PANREATWLSGYPTDSELYKLIASANAIIRNYAISKDTGFTVYKKNPIYKDDTTIAMRGT 420
   |||||||
362  PANREATWLSGYPTDSELYKLIASANAIIRNYAISKDTGFTVYKKNPIYKDDTTIAMRGT 421
   |||||||
421  DGSQIVTILSNKAGSDSYTILSGAGYTAGOQLLEVIGCTTIVYVGSNGNVPVPAAGLP 480
   |||||||
422  DGSQIVTILSNKAGSDSYTILSGAGYTAGOQLLEVIGCTTIVYVGSNGNVPVPAAGLP 481
   |||||||
481  RVLYPTEKLAGSKICSS 498
   |||||||
482  RVLYPTEKLAGSKICSS 499
   |||||||

```

RESULT 2
 alpha-amyase (EC 3.2.1.1) precursor - Aspergillus sp.

A:Accession: J50663
 C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 17-Mar-1999
 C:Accession: J50663
 R:Shibuya, I.; Tamura, G.; Ishikawa, T.; Hara, S.
 Biosci. Biotechnol. Biochem. 56, 174-179, 1992
 A:Title: Cloning of the alpha-amyase cDNA of Aspergillus shirousami and its expression
 A:Reference number: J50663; MUID:9223146
 A:Accession: J50663
 A:Molecule type: mRNA
 A:Residues: 1-499 <SH1>

C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycoen/starch degradation
 C:Superfamily: Aspergillus alpha-amyase; alpha-amyase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:1-21/Domain: signal sequence #status predicted <Sig>
 F:22-499/Product: alpha-amyase #status predicted <AMP>
 F:194-321/Domain: alpha-amyase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 99.7%; Score 2675; DB 2: Length 499;
 Best Local Similarity 99.8%; Pred. No. 5.6e-185;
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

1  MVAWMSLFLVGLQVAAAPALATPADMRSSOSIFLLTDFARTDSTTATCTADOKXCG 60
   |||||||
2  MVAWMSLFLVGLQVAAAPALATPADMRSSOSIFLLTDFARTDSTTATCTADOKXCG 61
   |||||||
61  TWGGIIDKLDYIOGMGFTAIWTPVTAQLPQTAYGDAYHGWOODIYSLNENGTADL 120
   |||||||
62  TWGGIIDKLDYIOGMGFTAIWTPVTAQLPQTAYGDAYHGWOODIYSLNENGTADL 121
   |||||||
121  KALSSALHERGMVLMVDVYANMHMGYDAGSSVDYSVFRPSSQDYFHPFCFIONYEOTO 180
   |||||||
122  KALSSALHERGMVLMVDVYANMHMGYDAGSSVDYSVFRPSSQDYFHPFCFIONYEOTO 181
   |||||||
181  VEDCWIGDNTVSLPDLDTTKDYVKNEMWDVGSVSNYSIDGLRIDYKHKQDFWPGYN 240
   |||||||
182  VEDCWIGDNTVSLPDLDTTKDYVKNEMWDVGSVSNYSIDGLRIDYKHKQDFWPGYN 241
   |||||||
241  KAAGVYCIGEVLDGDPATYCPYQNVMDGVLPYIYPLLNKFKSTSGSMDLLYNNINTVK 300
   |||||||
242  KAAGVYCIGEVLDGDPATYCPYQNVMDGVLPYIYPLLNKFKSTSGSMDLLYNNINTVK 301
   |||||||
301  SDCPSTLLGTFFVENHNDREFASYTNDIALAKNVAAFIILNDGIPPIYAGOEORHAGND 360
   |||||||
302  SDCPSTLLGTFFVENHNDREFASTINDIALAKNVAAFIILNDGIPPIYAGOEORHAGND 361
   |||||||
361  PANREATWLSGYPTDSELYKLIASANAIIRNYAISKDTGFTVYKKNPIYKDDTTIAMRGT 420
   |||||||
362  PANREATWLSGYPTDSELYKLIASANAIIRNYAISKDTGFTVYKKNPIYKDDTTIAMRGT 421
   |||||||
421  DGSQIVTILSNKAGSDSYTILSGAGYTAGOQLLEVIGCTTIVYVGSNGNVPVPAAGLP 480
   |||||||
422  DGSQIVTILSNKAGSDSYTILSGAGYTAGOQLLEVIGCTTIVYVGSNGNVPVPAAGLP 481
   |||||||
481  RVLYPTEKLAGSKICSS 498
   |||||||
482  RVLYPTEKLAGSKICSS 499
   |||||||

```

RESULT 3

alpha-amyase (EC 3.2.1.1) 3 precursor - Aspergillus oryzae

A:Accession: A33215
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
 C:Accession: S04549; A33215; A44713
 R:Wiersel, S.; Lachmund, A.; Wildhardt, G.; Rutkowski, E.
 Mol. Microbiol. 3, 3-14, 1989
 A:Title: Three alpha-amyase genes of Aspergillus oryzae exhibit identical intron-exo
 A:Reference number: S04548; MUID:89237897
 A:Accession: S04549
 A:Molecule type: DNA
 A:Residues: 1-499 <NR>
 A:Cross-references: EMBL:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-499 <NI2>
 A:Cross-references: GB:X12727; NID:g2454; PIDN:CAA31220.1; PID:g295922
 R:Genes: M.J.; Dove, M.J.; Seligy, V.L.

Gene 79, 107-117, 1989
A:Title: Aspergillus oryzae has two nearly identical Taka-amylose genes, each containing
A:Reference number: J50240; MUID:89378767
A:Accession: A44713
A:Molecule type: DNA
A:Residues: 1-499 <GEN>
A:Note: the authors refer to this as Isozyme I
R:Matsumura, Y.; Kusunoki, M.; Harada, W.; Kakudo, M.
J. Biochem. 95, 697-702, 1984
A:Title: Structure and possible catalytic residues of Taka-amylose A.
A:Reference number: A37454; MUID:84212370
A:Contents: annotation; X-ray crystallography, 3.0 angstroms
C:Comment: One atom of calcium per molecule is essential for activity.
C:Genetics:
A:Gene: amy3; Amy1
A:Introns: 56/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
C:Keywords: calcium; carbohydrate digestion; glycoprotein; hydrolase; metal
F:1-21/Domains: signal sequence #status predicted <SIG>
F:12-499/Product: alpha-amylose #3 status experimental <MAT>
F:194-321/Domains: alpha-amylose core homology <AMY>
F:51-59, 171-185, 261-304, 461-496/Disulfide bonds: #status experimental
F:142, 183, 196, 231/Binding site: calcium (Asn, Glu, Asp, His) #status predicted
F:218/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:231, 251, 318/Active site: His, Glu, Asp #status experimental

Query Match 99.6%; Score 2674; DB 1; Length 499;
Best Local Similarity 99.6%; Pred. No. 6, 6e-185;
Matches 496; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MWAMWSLFLYGLQVAAPALATPADWRSOSIFLLTRFARTDGTATCNADOKYCGG 60
DB 2 MWAMWSLFLYGLQVAAPALATPADWRSOSIFLLTRFARTDGTATCNADOKYCGG 61
QY 61 TWGGIIDLKLDYIQGMGFALWITPVTAQLPOTTAAGDAHGWOODIYSINENYGTADDL 120
DB 62 TWGGIIDLKLDYIQGMGFALWITPVTAQLPOTTAAGDAHGWOODIYSINENYGTADDL 121
QY 121 KALSALHERGMYLMVDVYANHMGYDAGSSVDYSVFKPSSODYFHPFCFIONYEDQTQ 180
DB 122 KALSALHERGMYLMVDVYANHMGYDAGSSVDYSVFKPSSODYFHPFCFIONYEDQTQ 181
QY 181 VEDCMLGDNVTSLPDLDTTKDYVKNEMVDWGSLSVNSYSDGLRIDTVKHVKQDFMPGYN 240
DB 182 VEDCMLGDNVTSLPDLDTTKDYVKNEMVDWGSLSVNSYSDGLRIDTVKHVKQDFMPGYN 241
QY 241 KAAGYVCIGEVLDGDPATTCPCYQNVMDGVNLNPIYPLLNFAFKSTSGSMDLYNMINTVK 300
DB 242 KAAGYVCIGEVLDGDPATTCPCYQNVMDGVNLNPIYPLLNFAFKSTSGSMDLYNMINTVK 301
QY 301 SDCPSTLLGTFVFNHNDPRFASVTNDIALAKNVAAFIINDGIPRIYAGQEQHYAGND 360
DB 302 SDCPSTLLGTFVFNHNDPRFASVTNDIALAKNVAAFIINDGIPRIYAGQEQHYAGND 361
QY 361 PANREATWLSGYPDSELYKLIAASNAIRNVAISKDGFVYTKNMPYIKDDTTIAPRKGT 420
DB 362 PANREATWLSGYPDSELYKLIAASNAIRNVAISKDGFVYTKNMPYIKDDTTIAPRKGT 421
QY 421 DGSQIVTLLSNKGASGDSYTLISGAGYTAGOQLTEVIGCTTVYVSGDGNVPVPMAGGLP 480
DB 422 DGSQIVTLLSNKGASGDSYTLISGAGYTAGOQLTEVIGCTTVYVSGDGNVPVPMAGGLP 481
QY 481 RVLYPTEKLAGSKICSS 498
DB 482 RVLYPTEKLAGSKICSS 499

RESULT 4
B48305

alpha-amylose (EC 3.2.1.1) B precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: B48305
R:Korman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylose genes from As
A:Reference number: A48305; MUID:90254827
A:Accession: B48305
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-499 <KOR>
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domains: alpha-amylose core homology <AMY>

Query Match 99.6%; Score 2674; DB 2; Length 499;
Best Local Similarity 99.6%; Pred. No. 6, 6e-185;
Matches 496; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MWAMWSLFLYGLQVAAPALATPADWRSOSIFLLTRFARTDGTATCNADOKYCGG 60
DB 2 MWAMWSLFLYGLQVAAPALATPADWRSOSIFLLTRFARTDGTATCNADOKYCGG 61
QY 61 TWGGIIDLKLDYIQGMGFALWITPVTAQLPOTTAAGDAHGWOODIYSINENYGTADDL 120
DB 62 TWGGIIDLKLDYIQGMGFALWITPVTAQLPOTTAAGDAHGWOODIYSINENYGTADDL 121
QY 121 KALSALHERGMYLMVDVYANHMGYDAGSSVDYSVFKPSSODYFHPFCFIONYEDQTQ 180
DB 122 KALSALHERGMYLMVDVYANHMGYDAGSSVDYSVFKPSSODYFHPFCFIONYEDQTQ 181
QY 181 VEDCMLGDNVTSLPDLDTTKDYVKNEMVDWGSLSVNSYSDGLRIDTVKHVKQDFMPGYN 240
DB 182 VEDCMLGDNVTSLPDLDTTKDYVKNEMVDWGSLSVNSYSDGLRIDTVKHVKQDFMPGYN 241
QY 241 KAAGYVCIGEVLDGDPATTCPCYQNVMDGVNLNPIYPLLNFAFKSTSGSMDLYNMINTVK 300
DB 242 KAAGYVCIGEVLDGDPATTCPCYQNVMDGVNLNPIYPLLNFAFKSTSGSMDLYNMINTVK 301
QY 301 SDCPSTLLGTFVFNHNDPRFASVTNDIALAKNVAAFIINDGIPRIYAGQEQHYAGND 360
DB 302 SDCPSTLLGTFVFNHNDPRFASVTNDIALAKNVAAFIINDGIPRIYAGQEQHYAGND 361
QY 361 PANREATWLSGYPDSELYKLIAASNAIRNVAISKDGFVYTKNMPYIKDDTTIAPRKGT 420
DB 362 PANREATWLSGYPDSELYKLIAASNAIRNVAISKDGFVYTKNMPYIKDDTTIAPRKGT 421
QY 421 DGSQIVTLLSNKGASGDSYTLISGAGYTAGOQLTEVIGCTTVYVSGDGNVPVPMAGGLP 480
DB 422 DGSQIVTLLSNKGASGDSYTLISGAGYTAGOQLTEVIGCTTVYVSGDGNVPVPMAGGLP 481
QY 481 RVLYPTEKLAGSKICSS 498
DB 482 RVLYPTEKLAGSKICSS 499

RESULT 5
JN0588
alpha-amylose (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: Taka-amylose A
C:Species: Aspergillus oryzae
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 14-Nov-1997
C:Accession: JN0588
R:Tsukagoshi, N.; Furukawa, M.; Nagaba, H.; Kiritla, N.; Tsuboi, A.; Uda, S.
Gene 84, 319-327, 1989
A:Title: Isolation of a cDNA encoding Aspergillus oryzae Taka-amylose A: evidence for
A:Reference number: JN0588; MUID:90128276
A:Accession: JN0588

A:Molecule type: mRNA
A:Residues: 1-499 <TSD>
C:Comment: The alpha amylases are encoded by multigene family.
C:Genetics:
A:Gene: Taa-GI
A:Introns: 56/1; 69/3; 108/2; 144/3; 221/1; 275/2; 324/2; 404/3
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <Sig>
F:22-499/Product: alpha-amylase #status predicted <AMT>
F:194-321/Domain: alpha-amylase core homology <AMT>
F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 99.6%; Score 2673; DB 2; Length 499;
Best Local Similarity 99.6%; Pred. No. 7.7e-185;
Matches 496; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MYAMWSLFLYGLQVAAPALATPADWRSQSYFLITDRFARTDGSSTATCNTADOKYCG 60
2 MYAMWSLFLYGLQVAAPALATPADWRSQSYFLITDRFARTDGSSTATCNTADOKYCG 61

61 TWOGIIDKLDYIOGMGFTAIWTPVTAQLPQTAYGDAHYGWODIYSLNENGTADDL 120
62 TWOGIIDKLDYIOGMGFTAIWTPVTAQLPQTAYGDAHYGWODIYSLNENGTADDL 121

121 KALSSALHERGMYLWVDVYANHMGYDAGSSVDYVFKPFSSQDYFHPFCFIQNYEDQTQ 180
122 KALSSALHERGMYLWVDVYANHMGYDAGSSVDYVFKPFSSQDYFHPFCFIQNYEDQTQ 181

181 VEDCWLGDNTVSLPDLDTTKDVKNKNEWYDWGSLVSNYSIDGLRIDYVKHYOKDFWPGYN 240
182 VEDCWLGDNTVSLPDLDTTKDVKNKNEWYDWGSLVSNYSIDGLRIDYVKHYOKDFWPGYN 241

241 KAAGYCTGEVLDDGPATYCPYONMDGVLNPIYPLNFAFKSTSGSMDDLKMYINTVK 300
242 KAAGYCTGEVLDDGPATYCPYONMDGVLNPIYPLNFAFKSTSGSMDDLKMYINTVK 301

301 SDPCDSTLLGTFEVENHNDNRFASYTNDIALAKNVAAFIILNDGIPITVAGQEQHYAGND 360
302 SDPCDSTLLGTFEVENHNDNRFASYTNDIALAKNVAAFIILNDGIPITVAGQEQHYAGND 361

361 PANREATWLSGYPTDSELYKLASNAIRNVAISKDTGVTYKMPYIKDDTTIAMRGT 420
362 PANREATWLSGYPTDSELYKLASNAIRNVAISKDTGVTYKMPYIKDDTTIAMRGT 421

421 DGSQIVTILSNKGASGDSYTLISGAGYTAGOQLFEVIGCTTVTVGSDGNVPVMAAGLP 480
422 DGSQIVTILSNKGASGDSYTLISGAGYTAGOQLFEVIGCTTVTVGSDGNVPVMAAGLP 481

481 RVLYPTEKLAGSKICSS 498
482 RVLYPTEKLAGSKICSS 499

RESULT 6
A:alpha-amylase (EC 3.2.1.1) A precursor - Aspergillus awamori
C:Species: Aspergillus awamori
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 14-Nov-1997
C:Accession: A48305
R:Kotman, D.R.; Bayliss, F.T.; Barnett, C.C.; Carmona, C.L.; Kodama, K.H.; Royer, T.J.;
Curr. Genet. 17, 203-212, 1990
A:Title: Cloning, characterization, and expression of two alpha-amylase genes from Aspergillus
A:Reference number: A48305; MUID:90254827
A:Accession: A48305
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-498 <KOR>
C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:194-321/Domain: alpha-amylase core homology <AMT>

Query Match 99.2%; Score 2662; DB 2; Length 498;
Best Local Similarity 99.6%; Pred. No. 4.8e-184;
Matches 493; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MYAMWSLFLYGLQVAAPALATPADWRSQSYFLITDRFARTDGSSTATCNTADOKYCG 60
2 MYAMWSLFLYGLQVAAPALATPADWRSQSYFLITDRFARTDGSSTATCNTADOKYCG 61

61 TWOGIIDKLDYIOGMGFTAIWTPVTAQLPQTAYGDAHYGWODIYSLNENGTADDL 120
62 TWOGIIDKLDYIOGMGFTAIWTPVTAQLPQTAYGDAHYGWODIYSLNENGTADDL 121

121 KALSSALHERGMYLWVDVYANHMGYDAGSSVDYVFKPFSSQDYFHPFCFIQNYEDQTQ 180
122 KALSSALHERGMYLWVDVYANHMGYDAGSSVDYVFKPFSSQDYFHPFCFIQNYEDQTQ 181

181 VEDCWLGDNTVSLPDLDTTKDVKNKNEWYDWGSLVSNYSIDGLRIDYVKHYOKDFWPGYN 240
182 VEDCWLGDNTVSLPDLDTTKDVKNKNEWYDWGSLVSNYSIDGLRIDYVKHYOKDFWPGYN 241

241 KAAGYCTGEVLDDGPATYCPYONMDGVLNPIYPLNFAFKSTSGSMDDLKMYINTVK 300
242 KAAGYCTGEVLDDGPATYCPYONMDGVLNPIYPLNFAFKSTSGSMDDLKMYINTVK 301

301 SDPCDSTLLGTFEVENHNDNRFASYTNDIALAKNVAAFIILNDGIPITVAGQEQHYAGND 360
302 SDPCDSTLLGTFEVENHNDNRFASYTNDIALAKNVAAFIILNDGIPITVAGQEQHYAGND 361

361 PANREATWLSGYPTDSELYKLASNAIRNVAISKDTGVTYKMPYIKDDTTIAMRGT 420
362 PANREATWLSGYPTDSELYKLASNAIRNVAISKDTGVTYKMPYIKDDTTIAMRGT 421

421 DGSQIVTILSNKGASGDSYTLISGAGYTAGOQLFEVIGCTTVTVGSDGNVPVMAAGLP 480
422 DGSQIVTILSNKGASGDSYTLISGAGYTAGOQLFEVIGCTTVTVGSDGNVPVMAAGLP 481

481 RVLYPTEKLAGSKIC 495
482 RVLYPTEKLAGSKIC 496

RESULT 7
JT0466
alpha-amylase (EC 3.2.1.1) precursor - Aspergillus oryzae
N:Alternate names: glycogenase; Taka-amylase A
C:Species: Aspergillus oryzae
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 10-Jul-1998
C:Accession: JT0466
R:Tada, S.; Iimura, Y.; Gomi, K.; Takahashi, K.; Hara, S.; Yoshizawa, K.
Agric. Biol. Chem. 53, 593-599, 1989
A:Title: Cloning and nucleotide sequence of the genomic Taka-amylase A gene of Aspergillus
A:Reference number: JT0466
A:Accession: JT0466
A:Molecule type: DNA
A:Residues: 1-499 <TAD>
C:Comment: See also PIR:J0201 and PIR:J0240.
C:Genetics: One atom of calcium per molecule is essential for activity.
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-21/Domain: signal sequence #status predicted <Sig>
F:22-499/Product: alpha-amylase #status predicted <AMT>

F:194-321/Domain: alpha-amylase core homology <AMY>
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:231,251,318/Active site: His, Glu, Asp #status predicted

Query Match 99.0%; Score 2656; DB 2; Length 499;
 Best Local Similarity 99.4%; Pred. No. 1,3e-183;
 Matches 495; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 MVAWMSILFLYGLQVAPALAAATPADMSOSITFLITDFARTDGTATCNATADOKYCGG 60
   |||
DB 2 MVAWMSILFLYGLQVAPALAAATPADMSOSITFLITDFARTDGTATCNATADOKYCGG 61
OY 61 TWOGIIDLKLDYIQGMGFTAWITPTAQLPOTAYGDAYHGYWOODIYSLNENYGTADDL 120
   |||
DB 62 TWOGIIDLKLDYIQGMGFTAWITPTAQLPOTAYGDAYHGYWOODIYSLNENYGTADDL 121
OY 121 KALSALHERGMYLWADVYANMGYDAGSSVDYVFPFPSSODYFHPFCFIONYEDQTO 180
   |||
DB 122 KALSALHERGMYLWADVYANMGYDAGSSVDYVFPFPSSODYFHPFCFIONYEDQTO 181
OY 181 VEDCWLIGDNTVSLPDLITTKDYVKNEMWDWGSIVSNYSIDGLRIDTVKHYOKDFWPGYN 240
   |||
DB 182 VECGWLIGDNTVSLPDLITTKDYVKNEMWDWGSIVSNYSIDGLRIDTVKHYOKDFWPGYN 241
OY 241 KAGVYCIIGEVLDGDPATYCPYQNMWDGVLNPIYPLLNAPKSGSMODLYNMINTVK 300
   |||
DB 242 KAGVYCIIGEVLDGDPATYCPYQNMWDGVLNPIYPLLNAPKSGSMODLYNMINTVK 301
OY 301 SDGPDSTLLGTFFVENHONRFASYNDIALAKNVAFTLLDGIPIIYAGOEQHYAGGND 360
   |||
DB 302 SDGPDSTLLGTFFVENHONRFASYNDIALAKNVAFTLLDGIPIIYAGOEQHYAGGND 361
OY 361 PANREATWISGYPTEDESEKLTASANAIRNVAISKDGFVYKKNMPTIKDOTTIAMRGT 420
   |||
DB 362 PANREATWISGYPTEDESEKLTASANAIRNVAISKDGFVYKKNMPTIKDOTTIAMRGT 421
OY 421 DGSQIVTILSNKAGSDSYTSLSGAGYTAGOQLTEVIGCTTVYVSGDNVPVMAGGLP 480
   |||
DB 422 DGSQIVTILSNKAGSDSYTSLSGAGYTAGOQLTEVIGCTTVYVSGDNVPVMAGGLP 481
OY 481 RVLYPTEFKLAGSKICSS 498
   |||
DB 482 RVLYPTEFKLAGSKICSS 499

```

A:Function: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:173-300/Domain: alpha-amylase core homology <AMY>
 F:197/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:210,230,297/Active site: His, Glu, Asp #status experimental

Query Match 93.3%; Score 2503.5; DB 2; Length 478;
 Best Local Similarity 97.9%; Pred. No. 1,2e-172;
 Matches 468; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

```

OY 21 ATPADMSOSITFLITDFARTDGTATCNATADOKYCGGWIIDKLDYIQGMGFTAI 80
   |||
DB 1 ATPADMSOSITFLITDFARTDGTATCNATADOKYCGGWIIDKLDYIQGMGFTAI 60
OY 81 WITPVTALPQTTAYGDAYHGYWOODIYSLNENYGTADDLKALSALHERGMYLWADVVA 140
   |||
DB 61 WITPVTALPQTTAYGDAYHGYWOODIYSLNENYGTADDLKALSALHERGMYLWADVVA 120
OY 141 NHMGYDAGSSVDYVFKFPSSODYFHPFCFIONYEDQTVEDCWLGNITYSLPDLITK 200
   |||
DB 121 NHMGYDAGSSVDYVFKFPSSODYFHPFCFIONYEDQTVEDCWLGNITYSLPDLITK 180
OY 201 DVKNEMWDWGSIVSNYSIDGLRIDTVKHYOKDFWPGYNKAGYCIIGEVLDGDPATYC 260
   |||
DB 181 DVKNEMWDWGSIVSNYSIDGLRIDTVKHYOKDFWPGYNKAGYCIIGEVLDGDPATYC 240
OY 261 PYQNMWDGVLNPIYPLLNAPKSGSMODLYNMINTVKSDDPSTLLGTFFVENHNDNR 320
   |||
DB 241 PYQNMWDGVLNPIYPLLNAPKSGSMODLYNMINTVKSDDPSTLLGTFFVENHNDNR 300
OY 321 FASTYNDIALAKNVAFTLLDGIPIIYAGOEQHYAGGNDPANREATWISGYPTESELYK 380
   |||
DB 301 FASTYNDIALAKNVAFTLLDGIPIIYAGOEQHYAGGNDPANREATWISGYPTESELYK 360
OY 361 LIASANAIRNVAISKDGFVYKKNMPTIKDOTTIAMRGTDSQIVTILSNKAGSDSYT 440
   |||
DB 361 LIASANAIRNVAISKDGFVYKKNMPTIKDOTTIAMRGTDSQIVTILSNKAGSDSYT 419
OY 441 LSLGAGYTAGOQLTEVIGCTTVYVSGDNVPVMAGGLPRVLYPTEFKLAGSKICSS 498
   |||
DB 420 LSLGAGYTAGOQLTEVIGCTTVYVSGDNVPVMAGGLPRVLYPTEFKLAGSKICSS 477

```

RESULT 9
 alpha-amylase (EC 3.2.1.1) - Aspergillus niger
 C:Species: Aspergillus niger
 C:Date: 10-Sep-1989 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A35282
 R:Boel, E.; Brady, L.; Brzozowski, A.M.; Derewenda, Z.; Dodson, G.G.; Jensen, V.J.; P
 Biochemistry 29, 6244-6249, 1990
 A:Title: Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-angstrom
 A:Reference number: A35282; PMID:91002514
 A:Accession: A35282
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A:Molecule type: mRNA
 A:Residues: 1-484 <BOE>
 C:Function:
 A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
 A:Pathway: glycogen/starch degradation
 C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
 C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
 F:173-300/Domain: alpha-amylase core homology <AMY>

Query Match 66.3%; Score 1780; DB 1; Length 484;
 Best Local Similarity 67.2%; Pred. No. 1,6e-120;
 Matches 319; Conservative 63; Mismatches 93; Indels 0; Gaps 0;

```

OY 24 ADMRSOSITFLITDFARTDGTATCNATADOKYCGGWIIDKLDYIQGMGFTAIWIT 83
   |||
DB 4 ADMRSOSITFLITDFARTDGTATCNATADOKYCGGWIIDKLDYIQGMGFTAIWIS 63
OY 84 PVTALPQTTAYGDAYHGYWOODIYSLNENYGTADDLKALSALHERGMYLWADVVA 143
   |||
DB 64 PVTALPQTTAYGDAYHGYWOODIYSLNENYGTADDLKALSALHERGMYLWADVVA 123
OY 144 GTDAGSSVDYVFKFPSSODYFHPFCFIONYEDQTVEDCWLGNITYSLPDLITTKDY 203
   |||

```

```

Db 124 GYAGNGNDVYSVFEDPSSSYFHPYCLITDMNLTMVQDCGFTIYSLPDLNTETFAV 183
OY 204 KNEVDWVGSLSVNSIDGLRIDTVKHKVQKDFWPYKNAAGYVIGEDLDGPATTCPYQ 263
Db 184 RITWDMVADVSNYSVDGLRIDSVLEVPDFPFQYQEAAGYVCGEVDNCPALDCPYQ 243
OY 264 NMDGVLPYPIYPLINAFKSTSGMDLYNMINTVKSQCDPSTLLGTFFVENHNDPRAS 323
Db 244 KYLDGVLNPTFWOLLYAFESSSGSISNLYNMKISVASCSDPTLLGNTEHNDPRAS 303
OY 324 YNMDILAKNVAFFILNDGIPITVYAGEQHYAGGNDPANCPEPTSELYKILIA 383
Db 304 YTSDSQANNNVSYIFLSGIPVYAGEQHYSGKVPYINRATMLSGDTSAELXTWIA 363
OY 384 SANAINVAISKDTGFVYKKNPIYKDDTTIAMRKGTDSQIVTILSNKASGDSYTLSL 443
Db 364 TTAIRKILIASDSAYITVANDAFYDSNTIAMRKGTSGSOVITVLSNKGSSGSTYTL 423
OY 444 SGAGYTAGOOLTEVIGCTTVYVSGDGNVPVMAAGLPRLVLPTEKLAGSKISS 498
Db 424 SSGSYTSGTKLIEAVYCTSVYDSSGDIPEPMASGLPRVLLPVSVDSSSLCGGS 478

RESULT 10
JC4510
alpha-amylase (EC 3.2.1.1) precursor - yeast (Lipomyces kononenkoae)
N:Alternate names: LKA1 protein; raw starch-degrading amylase
N:Contains: alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41)
C:Species: Lipomyces kononenkoae
C>Date: 07-Feb-1996 #sequence_revision 23-Aug-1996 #text_change 18-Jun-1999
C:Accession: J04510; M0ID:96105202
Gene 166, 65-71, 1995
A:Title: Cloning, sequence analysis and expression in yeasts of a cDNA containing a Lipo
A:Reference number: J04510; M0ID:96105202
A:Accession: J04510
A:Molecule type: mRNA
A:Residues: 1-624 <STP>
A:Cross-references: GB:030376; NID:q1173536; PIDN:AAC49622.1; PID:q1173537
A:Experimental source: strain IGC4052B
A:Molecule type: Protein
A:Residues: 29-44 <ST2>
A:Experimental source: IGC4052B
C:Genetics:
A:Gene: LKA1
C:Function:
A:Description: hydrolyzes raw starch, degrading both alpha-1,4 and alpha-1,6 linkages
A:Pathway: glycogen/starch degradation
C:Superfamily: Lipomyces alpha-amylase; alpha-amylase core homology; glucoamylase starch
C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide
C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide
F:28/Domain: signal sequence #status predicted <SIC>
F:29-624/Product: alpha-amylase #status predicted <MAT>
F:48-141/Domain: glucoamylase starch-binding domain homology <SBD>
F:320-447/Domain: alpha-amylase core homology <AMY>
F:177-185,297-311,387-430,587-622/Disulfide bonds: #status predicted
F:304,344/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:337,377,444/Active site: His, Glu, Asp #status predicted

```

```

Query Match 60.7%; Score 1629; DB 1; Length 624;
Best Local Similarity 60.6%; Pred. No. 1.8e-109;
Matches 292; Conservative 77; Mismatches 113; Indels 0; Gaps 0;

OY 15 AAPALATPADWRSOSITFLITDFARPDTGTATCNTAOKYGGTMOGTLDTYIOG 74
Db 142 SASVPTGTAAMWGRSIVQVYVTDRAFDGSIITSCDVTDVYGGSGYRGLTNMLDYIOG 201
OY 75 MGFATWTPVTAOLPOTTAAGDAYHGYWODIYSLNENYATADLALSSALHERGMYL 134
Db 202 MGFATWTPVTAOLPOTTAAGDAYHGYWODIYSLNENYATADLALSSALHERGMYL 261
OY 135 MVDVANHMGYDGAGSSVDYSVFRFSSQDYFHPFCFIQNTYEDQTVQEDCWLGNVSLP 194

Db 262 MVDIVVHHPAFSGNHADVDSSEYFPYSSQDYFHSFCMIDTYSNQVNEECWLDSDSVLY 321
OY 195 DLDFTTKDVKNWYDWVGLSVNSIDGLRIDTVKHKVQKDFWPYKNAAGYVIGEDLD 254
Db 322 DVNTQDLTVSEYOSVWKOILIANSIDGLRIDTVKHKVQKDFWPYKNAAGYVIGEDLD 381
OY 255 DPATCPYQNMGVNLNPIYPLINAFKSTSGMDLYNMINTVKSQCDPSTLLGTFFVE 314
Db 382 DPSTCPYQNMGVNLNPIYPLINAFKSTSGMDLYNMINTVKSQCDPSTLLGTFFVE 441
OY 315 NHDNPFASVYNDITALKNVAFFILNDGIPITVYAGEQHYAGGNDPANCPEPTSELY 374
Db 442 NQDNPRFPSYTSDESLIKNIAFTILSDGIPITVYAGEQHYAGGNDPANCPEPTSELY 501
OY 375 DSELYKILIASANAINVAISKDTGFVYKKNPIYKDDTTIAMRKGTDSQIVTILSNK 434
Db 502 TSTFEYVYASLNOIRNAIYIDTVYQWVYISDSTTAMRKGTGFTGNILTVLSMGS 561
OY 435 SGDSYTLISGAGYTAGOOLTEVIGCTTVYVSGDGNVPVMAAGLPRLVLPTEKLAGSK 494
Db 562 SSGSYTTLTSLNKGSTVYVSGDGNVPVMAAGLPRLVLPTEKLAGSK 621
OY 495 CS 496
Db 622 CS 623

RESULT 11
S33921
alpha-amylase (EC 3.2.1.1) SWA2 precursor - yeast (Schwanniomyces occidentalis)
N:Alternate names: alpha-1,4 glucanohydrolase
N:Contains: Schwanniomyces occidentalis
C:Species: Schwanniomyces occidentalis
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S33921
R:Claros, M.G.; Abarea, D.; Fernandez-Lobato, M.; Jimenez, A.
Curr. Genet. 24, 75-83, 1993
A:Title: Molecular structure of the SWA2 gene encoding an AM1-related alpha-amylase
A:Reference number: S33921; M0ID:93365041
A:Accession: S33921
A:Molecule type: DNA
A:Residues: 1-507 <CTA>
A:Cross-references: EMBL:X73497; NID:9336561; PIDN:CAA51912.1; PID:g3396562
C:Genetics:
A:Gene: SWA2
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: Aspergillus alpha-amylase; alpha-amylase core homology
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-507/Product: alpha-amylase #status predicted <MAT>
F:205-332/Domain: alpha-amylase core homology <AMY>
F:134,223/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

```

Query Match 56.9%; Score 1527; DB 2; Length 507;
Best Local Similarity 57.7%; Pred. No. 3e-102;
Matches 286; Conservative 74; Mismatches 130; Indels 6; Gaps 3;

```

```

OY 1 MYAAMSLPFLXGQVAPALATPADWRSOSITFLITDFARPDTGTATCNTAOKYCG 60
Db 17 LVASKRPIFLSKRDAGSAA--WRSSTIYQVYVTDRAFDGSIITSCDVTDVYGG 72
OY 61 TMOGITDKLDYIOGMGFTAWITPVAOLPOTTAAGDAYHGYWODIYSLNENYATADL 120
Db 73 TPOGITDKLDYIOGKFTAWITSPVEQIPDDTGYGAYHGYWODIYAINNFTADL 132
OY 121 KALSSALHERGMYLWVDVANHMGYDGAGSSVDYSVFRFSSQDYFHPFCFIQNTYED 180
Db 133 KNLNSELHRKMKLWVDIYTNHAYNAGSSVAYSNVFNQOYSYFHDCLITNDDQTN 192
OY 181 VEDCWLGNVSLPDLDTTKDVKNWYDWVGLSVNSIDGLRIDTVKHKVQKDFWPYGN 240

```

```

Db 193 VEDWEDBNVTYSLPDLTETEDSDVSSIFNLMVAVELVSNYSIDLRLIDSAKHVESPEPSFQ 252
OY 241 KAAGYVCIGEVLDGDPATYCPYQNVMDGYLVNPIYVPLLNARSTSGSMDDLNNYNNVK 300
Db 253 SAAGVYLIGEVLDGDPATYCPYQNVMDGYLVNPIYVPLLNARSTSGSMDDLNNYNNVK 312
OY 301 SDGPDSTLLGTEVENHNDNPRFASYNIDIALAKNVAAPILLNDGIPPIYAGQEQHYAGND 360
Db 313 SDCKDITLLGFENHNDQPRLPSTSDSALIKNAIFNLMDSGIDPIIYAGQEQHYAGND 372
OY 361 PANREATWLSGYPDSLEYKLTIASANAIRNYAISKDTGVTKKNNPIYKDDTTIMRGT 420
Db 373 PNNREALMLSGYSTNSGTYKLISVNDIRNOAIYKDSKTTTWSVDYLASGHVIALORGA 432
OY 421 DGSQIVTILSKGASDSTYLSLGGAGYTAGOQLTEVICCTVTYVSGDGNVPYMAAGLIP 480
OY 433 DDQRTVSVFNLLGSSG-SQHYTFSS-TRYSQGEKVVDVLTCQTSVANSDBTLTVSISGAP 490
OY 481 RVLYPTEKLASKTICS 496
Db 491 RYAPASTLIANSIGICN 506

```

RESULT 12

S06115

alpha-amylose (EC 3.2.1.1) AMY1 precursor - yeast (Schwanniomycetes occidentalis)

C:Species: Schwanniomycetes occidentalis

C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 22-Jun-1999

C:Accession: S06115

R:Strasser, A.W.M.; Selk, R.; Dohmen, R.J.; Niermann, T.; Bielefeld, M.; Seeboth, P.; Tu

Eur. J. Biochem. 184, 699-706, 1989

A:Title: Analysis of the alpha-amylose gene of Schwanniomycetes occidentalis and the secre

A:Reference number: S06115, NCID:90032659

A:Accession: S06115

A:Molecule type: DNA

A:Residues: 1-512 <STR>

A:Cross-references: EMBL:X16040; NID:g4882; PIDN:CAA34162.1; PID:g4883

C:Genetics:

A:Gene: AMY1

C:Function: A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycoegen/starch degradation

C:Keywords: Aspergillus alpha-amylose; alpha-amylose core homology

C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation

C:Keywords: 1-25/Domain: signal sequence #status predicted <SIS>

C:Keywords: 26-512/Product: alpha-amylose #status predicted <MAT>

C:Keywords: 209-336/Domain: alpha-amylose core homology <AMY>

Query Match 54.0%; Score 1449; DB 2; Length 512;

Best Local Similarity 56.4%; Pred. No. 1.3e-96;

Matches 269; Conservative 74; Mismatches 132; Indels 2; Gaps 2;

```

OY 21 ATPAD-WRSOSIYFLTLDRFARTGTTATCNTADQKYGCGTMOGIIIDKLDIYOGMGFTA 79
Db 36 SSSADKWKQDSITYOIVDRFARSDGTTADCLVSDKRYCGSGTKGIIDKLDIYOGMGFTA 95
OY 80 IWTTPVTAQLPQTTAYGDAYHGYWQODIYSLNENYGTADLKLASSALHERGMYLMDVY 139
Db 96 IWTTPVTAQLPQTTAYGDAYHGYWQODIYSLNENYGTADLKLASSALHERGMYLMDVY 155
OY 140 ANHMGYDAGSSVDYVFKPPSSODYFHPRCFIONYEDQTOVEDCWLGNVTYSLPDLSTE 199
Db 156 YNHIAWNGDSSVDISSLTPFNQOSTYFHDCLITNTNDQTNVEDCWEGDTEVSLPDLSTE 215
OY 200 KDYVKNEMYDMWGLSVNYSIDGLRIDTVKHOKDFWPYGNKAGYVCIGEVLDGDPATY 259
Db 216 DNEVIGVFOQVWSDVQVYNSIDGLRIDSAKHVDTASLTKFEDASGYVNLGEVYQGPPT 275
OY 260 CPYQNVMDGYLVNPIYVPLLNARSTSGSMDDLNNYNNVKSDGPDSTLLGTEVENHNDP 319
Db 276 CPYQNVMDGYLVNPIYVPLLNARSTSGSMDDLNNYNNVKSDGPDSTLLGTEVENHNDP 335

```

```

OY 320 RPSATYNDIALAKNVAAPILLNDGIPPIYAGQEQHYAGNDPANNREATWLSGYPDSLEY 379
Db 336 RPSATYNDIALAKNVAAPILLNDGIPPIYAGQEQHYAGNDPANNREATWLSGYPDSLEY 395
OY 380 KLIASANAIRNYAISKDTGVTKKNNPIYKDDTTIMRGTDSQIVTILSKGASDSTYLS 439
Db 396 ELISKNDIRNOAIYKDSKTTTWSVDYLASGHVIALORGA 454
OY 440 TSLSGAGYTAGOQLTEVICCTVTYVSGDGNVPYMAAGLIPRVLYPTEKLASKTICS 496
Db 455 DTTVSNVGTSSGDKYIIDISONSVLADSGSLYSISGMPYAYARSSVLSGSGICN 511

```

RESULT 13

S23355

alpha-amylose (EC 3.2.1.1) precursor - yeast (Schwanniomycetes occidentalis)

C:Species: Schwanniomycetes occidentalis

C:Date: 22-Nov-1993 #sequence_revision 13-Mar-1997 #text_change 22-Jun-1999

C:Accession: S23355

R:Mu, F.M.; Wang, T.T.; Hsu, W.H.

FEMS Microbiol. Lett. 82, 313-318, 1991

A:Title: The nucleotide sequence of Schwanniomycetes occidentalis alpha-amylose gene.

A:Reference number: S23355

A:Accession: S23355

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-512 <MUR>

A:Cross-references: EMBL:X62079; NID:g4880; PIDN:CAA43995.1; PID:g4881

A:Note: The authors translated the codon AGA for residue 21 as Pro, CTT for residue 6

C:Superfamily: Aspergillus alpha-amylose; alpha-amylose core homology

C:Keywords: glycoprotein; glycosidase; hydrolase

F:209-336/Domain: alpha-amylose core homology <AMY>

Query Match 53.9%; Score 1445.5; DB 2; Length 512;

Best Local Similarity 56.5%; Pred. No. 2.2e-96;

Matches 266; Conservative 74; Mismatches 130; Indels 1; Gaps 1;

```

OY 26 WRSOSIYFLTLDRFARTGTTATCNTADQKYGCGTMOGIIIDKLDIYOGMGFTA 85
Db 42 WKQDSITYOIVDRFARSDGTTADCLVSDKRYCGSGTKGIIDKLDIYOGMGFTA 101
OY 86 TQQLPQTTAYGDAYHGYWQODIYSLNENYGTADLKLASSALHERGMYLMDVYANHMGY 145
Db 102 VEQIPNTAYGDAYHGYWQODIYSLNENYGTADLKLASSALHERGMYLMDVYANHMGY 161
OY 146 DAGSSVDYVFKPPSSODYFHPRCFIONYEDQTOVEDCWLGNVTYSLPDLSTEDEYVIG 205
Db 162 NGDSSVDYVFKPPSSODYFHPRCFIONYEDQTOVEDCWLGNVTYSLPDLSTEDEYVIG 221
OY 206 EYTDWVGLSVNYSIDGLRIDTVKHOKDFWPYGNKAGYVCIGEVLDGDPATYCPYQNV 265
Db 222 VEQVWSDVQVYNSIDGLRIDSAKHVDTASLTKFEDASGYVNLGEVYQGPPT 281
OY 266 MGVVNYPIYVPLLNARSTSGSMDDLNNYNNVKSDGPDSTLLGTEVENHNDP 325
Db 282 MGVVNYPIYVPLLNARSTSGSMDDLNNYNNVKSDGPDSTLLGTEVENHNDP 341
OY 326 NDIALAKNVAAPILLNDGIPPIYAGQEQHYAGNDPANNREATWLSGYPDSLEYKLASA 385
Db 342 SDTSLIKNAIFNLMDSGIDPIIYAGQEQHYAGNDPANNREATWLSGYPDSLEYKLASA 401
OY 386 NAIRNYAISKDTGVTKKNNPIYKDDTTIMRGTDSQIVTILSKGASDSTYLS 445
Db 402 NAIRNYAISKDTGVTKKNNPIYKDDTTIMRGTDSQIVTILSKGASDSTYLS 460
OY 446 AGYTAGOQLTEVICCTVTYVSGDGNVPYMAAGLIPRVLYPTEKLASKTICS 496
Db 461 TGVSGDKYIIDISONSVLADSGSLYSISGMPYAYARSSVLSGSGICN 511

```

RESULT 14

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:53:35 ; Search time 125.81 Seconds
(Without alignments)
684.774 Million cell updates/sec

Title: US-09-710-339-2
Perfect score: 2684
Sequence: 1 MVAWMSLFYGLQVAPALAA.....LPRVLVPTREKLAKSCSS 498

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2674	99.6	499	3	096TH4	096TH4 aspergillus
2	1826	68.0	490	3	09UV07	09UV07 emeticella
3	1810	67.4	640	3	013236	013236 aspergillus
4	1725.5	64.3	623	3	09UV09	09UV09 emeticella
5	1629	60.7	624	3	001117	001117 lipomyces k
6	1527	56.9	607	3	008806	008806 debaryomyces
7	1264	47.1	631	3	092394	092394 cryptococcus
8	963.5	35.9	625	3	074922	074922 schizosacch
9	920.5	34.3	491	3	013956	013956 schizosacch
10	898.5	33.5	564	3	09Y759	09Y759 schizosacch
11	737	27.5	482	2	060051	060051 thermocactin
12	653.5	24.3	774	3	042918	042918 schizosacch
13	493.5	18.4	323	3	096WR6	096WR6 schizosacch
14	453	16.9	1798	2	09K211	09K211 streptomyces
15	447.5	16.7	713	2	09F5M3	09F5M3 bacillus ci
16	428	15.9	725	2	059239	059239 bacillus sp

17	418.5	15.6	739	1	09UWN2	09UWN2 thermococcus
18	418	15.6	714	2	052766	052766 paenibacill
19	410	15.3	483	16	09RUP8	09RUP8 deinococcus
20	408	15.2	704	2	082984	082984 bacillus sp
21	404.5	15.1	1358	2	P70983	P70983 bacillus sp
22	403.5	15.0	692	2	030565	030565 bacillus br
23	399	14.9	711	16	0992R3	0992R3 streptococ
24	393.5	14.7	1021	2	09K5L6	09K5L6 actinoplan
25	392.5	14.6	711	2	092A00	092A00 bacillus st
26	378	14.1	588	2	065007	065007 thermus sp.
27	376.5	14.0	1104	2	006812	006812 micrococcus
28	373.5	13.9	580	16	09A609	09A609 caulobacter
29	369	13.7	588	2	09A1V2	09A1V2 bacillus st
30	368	13.7	588	2	093R22	093R22 bacillus st
31	367.5	13.7	724	2	09K5L5	09K5L5 actinoplan
32	363	13.5	690	16	09K186	09K186 vibrio chol
33	343.5	12.8	677	2	09KRH1	09KRH1 klebsiella
34	336.5	12.5	578	16	09K856	09K856 bacillus ha
35	333	12.4	1684	2	003658	003658 unidentified
36	332	12.4	589	2	09R9H8	09R9H8 bacillus su
37	332	12.4	589	16	006988	006988 bacillus su
38	329.5	12.3	958	16	09KFR4	09KFR4 bacillus ha
39	328.5	12.2	567	16	0992B2	0992B2 streptococ
40	324	12.1	558	2	059226	059226 bacillus sp
41	320	11.9	587	2	082982	082982 bacillus sp
42	320	11.9	591	16	0929P3	0929P3 listeria in
43	315.5	11.8	590	2	045490	045490 bacillus st
44	307.5	11.5	526	2	060102	060102 xanthomonas
45	305	11.4	955	2	045574	045574 bacillus sp

ALIGNMENTS

RESULT 1
ID Q96TH4 PRELIMINARY; PRT; 499 AA.
AC Q96TH4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TAKA-AMYLASE A (EC 3.2.1.1) (ALPHA-AMYLASE).
GN AMYA OR AMY1.
OS Aspergillus oryzae, and
OS Aspergillus flavus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062, 5059;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=A.oryzae; STRAIN=RTB40;
RX MEDLINE=20289310; PubMed=10830498;
RA Gomi K., Akeno T., Minehoki T., Ozeki K., Kumagai C., Okazaki N., Iimura Y.;
RT "Molecular cloning and characterization of a transcriptional activator gene, amyR, involved in the amylolytic gene expression in Aspergillus oryzae";
RT Biosci. Biotechnol. Biochem. 64:816-827(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=A.flavus; STRAIN=86-10D;
RA Fakhouy A.M., Woloshuk C.P.;
RT "Amyl, the alpha-amylase gene of Aspergillus flavus: Involvement in aflatoxin biosynthesis in maize kernels";
RT Phytopathology 89:908-914(1999).
DR EMBL; AF021876; BAA95703.1; -
DR EMBL; AF139925; AAF14264.1; -
SQ SEQUENCE 499 AA; 54804 MW; EEF42ADA71D20DA9 CRC64;

Query Match 99.6%; Score 2674; DB 3; Length 499;
Best Local Similarity 99.6%; Pred. No. 2.3e-174;
Matches 496; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 MYAMMSLFYGLQVAPALAAATPADMRSSQSYFLTDFRFTDGSSTATCMTADQKCYG 60
    |||||
DB 2 MYAMMSLFYGLQVAPALAAATPADMRSSQSYFLTDFRFTDGSSTATCMTADQKCYG 61
OY 61 TWOGIIDKLDYIQGMGFTAIWITPVTAOLPQTTAYGDAYHGYWOODIYSLNENGTADL 120
    |||||
DB 62 TWOGIIDKLDYIQGMGFTAIWITPVTAOLPQTTAYGDAYHGYWOODIYSLNENGTADL 121
OY 121 KALSSALHRGMVLMVDVYANHMGYDAGSSVDYVFKPFSSODYFHPCLQNTEDTO 180
    |||||
DB 122 KALSSALHRGMVLMVDVYANHMGYDAGSSVDYVFKPFSSODYFHPCLQNTEDTO 181
OY 181 VEDCMLGDMTVSLPDLDTKDVKNEMVWVSGSLVSNYSIDGLRIDTVKHQKDPWPGYN 240
    |||||
DB 182 VEDCMLGDMTVSLPDLDTKDVKNEMVWVSGSLVSNYSIDGLRIDTVKHQKDPWPGYN 241
OY 241 KAAGVYCIGEVLDDGPAYTCYQONWMDGVLPYIYPLLNAPKSTSGSMDDLYNMINTVK 300
    |||||
DB 242 KAAGVYCIGEVLDDGPAYTCYQONWMDGVLPYIYPLLNAPKSTSGSMDDLYNMINTVK 301
OY 301 SDCPDSLTLLGTVEVNDNFRFASYNDDIALAKNVAFFILLNDGIPLIYAGQOHYAGND 361
    |||||
DB 302 SDCPDSLTLLGTVEVNDNFRFASYNDDIALAKNVAFFILLNDGIPLIYAGQOHYAGND 361
OY 361 PANREATWLSGYPDTSELYKLIASANAIRNYAISKDTGFVTKNMPYIKDDTTIAMRKGT 420
    |||||
DB 362 PANREATWLSGYPDTSELYKLIASANAIRNYAISKDTGFVTKNMPYIKDDTTIAMRKGT 421
OY 421 DGSQIVTILSNKGASGDSYTLISLGAGYTAAGQOLTEVIGCTTVTVGSDGNVPMAGLIP 480
    |||||
DB 422 DGSQIVTILSNKGASGDSYTLISLGAGYTAAGQOLTEVIGCTTVTVGSDGNVPMAGLIP 481
OY 481 RVLYPTEKLAGSKICSS 498
    |||||
DB 482 RVLYPTEKLAGSKICSS 499

RESULT 2
OY 09UV07 PRELIMINARY: PRT: 490 AA.
AC 09UV07;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE ALPHA-AMYLASE AMYA.
GN AMYA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eucristales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.:
RA Boase N.A., Murphy R.M., Kelly J.M.;
RT "An amylase cluster in Aspergillus nidulans."
RT Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF208225; AAF17103.1; -.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha_amylase.
DR Pfam: PF00128; alpha-amylase; 1.
DR Pfam: PF00128; alpha-amylase; 1.
DR PRINTS: PR00110; ALPHAAMYLASE.
SQ SEQUENCE 490 AA; 54249 MW; A891C4CAEAB5305 CRC64;

```

Query Match 68.0%; Score 1826; DB 3; Length 490;
 Best Local Similarity 69.6%; Pred. No. 1.5e-116;
 Matches 337; Conservative 53; Mismatches 92; Indels 2; Gaps 2;

```

OY 12 LOVAALALATPADMRSSQSYFLTDFRFTDGSSTATCMTADQKCGGTWOGIIDKLDY 71
    |||||
DB 7 LSPFAATVLAATPAEMRSQSYFLTDFRFTDGSSTATCMTADQKCGGTWOGIIDKLDY 65

```

```

OY 72 IOGMGFTAIWITPVTAOLPQTTAYGDAYHGYWOODIYSLNENGTADDLKALSSALHERG 131
    |||||
DB 66 IOGMGFTAIWITPVTAOLPQTTAYGDAYHGYWOODIYALNPHRGTODDLKALSSALHERG 124
OY 132 MYLWADVAAHMGYDAGSSVDYVFKPFSSODYFHPFCIQYEDQOYQVEDCMLDNTY 191
    |||||
DB 125 MYLWADVAAHMGYDAGSSVDYVFKPFSSODYFHPFCIQYEDQOYQVEDCMLDNTY 184
OY 192 SLPDLDTKDVKNEMVWVSGSLVSNYSIDGLRIDTVKHQKDPWPGYNKAAGVYCI 251
    |||||
DB 185 SLPDLDTKDVKNEMVWVSGSLVSNYSIDGLRIDTVKHQKDPWPGYNKAAGVYCI 244
OY 252 LDGDPAYTCYQONWMDGVLPYIYPLLNAPKSTSGSMDDLYNMINTVKSDCPDSTLLG 311
    |||||
DB 245 LDGDPAYTCYQONWMDGVLPYIYPLLNAPKSTSGSMDDLYNMINTVKSDCPDSTLLG 304
OY 312 FVENHNDNFRFASYNDDIALAKNVAFFILLNDGIPLIYAGQOHYAGNDPANREATWLSG 371
    |||||
DB 305 FVENHNDNFRFASYNDDIALAKNVAFFILLNDGIPLIYAGQOHYAGNDPANREATWLSG 364
OY 372 YPTDSELYKLIASANAIRNYAISKDTGFVTKNMPYIKDDTTIAMRKGTGSOIVTILSN 431
    |||||
DB 365 YPTDSELYKLIASANAIRNYAISKDTGFVTKNMPYIKDDTTIAMRKGTGSOIVTILSN 424
OY 432 KGASGDSYTLISLGAGYTAAGQOLTEVIGCTTVTVGSDGNVPMAGLIPRVLYPTEKLAG 491
    |||||
DB 425 AGADAGSSYTVSVPNTGTTAGAAVTEIYTCEDITVSGSGEVSVEMSEGLPRVLYPTEKLAG 484
OY 492 SKIC 495
    |||||
DB 485 SKIC 488

RESULT 3
OY 013296 PRELIMINARY: PRT: 640 AA.
AC 013296;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE ACID-STABLE ALPHA-AMYLASE.
OS Aspergillus kawachii.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eucristales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=40384;
RN [1]
RP SEQUENCE FROM N.A.
RA kaneko A.;
RT Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA kaneko A., Sudo S., Sakamoto Y., Tamura G., Ishikawa T., Ohba T.;
RT "Molecular cloning and determination of the nucleotide sequence of a
RT gene encoding an acid-stable alpha-amylase from Aspergillus-
RT kawachii."
RT J. Ferment. Bioeng. 81:292-298(1996).
DR EMBL: AB008370; BAA2293.1; -.
DR HSSP: P56271; 2AAA.
DR InterPro: IPR000461; Alpha_amylase.
DR InterPro: IPR002044; CBD_4.
DR Pfam: PF00128; alpha-amylase; 1.
DR Pfam: PF00128; alpha-amylase; 1.
DR PRODOM: PD001568; CBD_4; 1.
SQ SEQUENCE 640 AA; 63507 MW; 676BF9D0236720DD CRC64;

```

Query Match 67.4%; Score 1810; DB 3; Length 640;
 Best Local Similarity 67.6%; Pred. No. 2.6e-115;
 Matches 325; Conservative 66; Mismatches 90; Indels 0; Gaps 0;

```

OY 18 ALAATPADMRSSQSYFLTDFRFTDGSSTATCMTADQKCGGTWOGIIDKLDYIQGMG 77
    |||||
DB 19 ALGSAAEWRQSYFLTDFRFTDGSSTATCMTADQKCGGTWOGIIDKLDYIQGMG 78

```

OY	78	TATWTPPTAQLPQTAAAGDAVHGWODIYSLENENGTADDLKALSALHERGMTAVD	137
OY	79	TATWTPPTAQLPQTASGDEAVHGWOQRIYYVNSNFGTADDLKSALDHARGMTAVD	138
OY	138	VVANMGVDGAGSVSDYVFVKPFSSODEFHPFCFTONYEDOTOVEDCMLGNMVSPLDL	197
Db	139	VYPNMGMAGANGNDVDYVFDFPPDSSSTFHRYCLITDMDNLTMYQDCWEGBDTIVSLFDLN	198
OY	198	TTTKDVVNKEWDWGSGLSVSNYSIDGLRIDTVKHVQDKFWPGYNKAAGVCIGEVLGDPA	257
Db	199	TTEETAVRTIWDVDWADVLSNYSVGCLRIDSVEEVEBPFFPGCAAGVYCGAEVDCNPA	258
OY	258	YTCPYQNMDGVNLNPITYPIPLNAFKSISGSMDOLYMNINTVSKSDCEPDSTLLTGFEVNH	317
Db	259	LDCPEOKLDELGVNLPIYWOLLVFAESSGSISIMLVNIKRSVASDCDPRTLGNFIENHD	318
OY	318	NPRFASYNDIALAKNVAAFIITLDNGJIIYAOGEOHYAGCNDPANSEAFMWLSGPYDSE	377
Db	319	NPRFASYSDXSQAKNVLSYIFLSDGIPIYVAGEBOHYSGSDVYVNEATWLSGYDTSAE	378
OY	378	LYKLIAASANAIRNVAISKDTGFVTYKKMPPIYKDDTTIAMRKGTDSOIVTLILSNKGASD	437
Db	379	LYTWIATINATKRILKLAISADSYIITYKNDPIYTDSENTIAMRGTSQSIOITIVLSMKSGSS	438
OY	438	SYTTLSSAGCTAGAOOLEVIGCTTIVYGSGNANPVPMAGGLPRVLYPTREKLASKRICSS	497
Db	439	SYTTLSSGSGYTSCKLTLEAYCTSVTVSDNSMGDIIPVMASGLPRVLLPASVVDSSSLCGG	498
OY	498	S 498	
Db	499	S 499	
RESULT	4		
O9UV09	PRELIMINARY;	PRT:	623 AA.
AC	O9UV09-		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)		
DE	ALPHA-AMYLASE.		
CN	AMYB.		
OS	Emericella nidulans (Aspergillus nidulans).		
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;		
OC	Eurotiales; Trichocomaceae; Emericella.		
OX	NCBI_Taxid=5072;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Boase N.A., Murphy R.L., Kelly J.M.;		
RT	"Characterisation and regulation of an alpha-amylase gene in		
RT	Aspergillus nidulans.";		
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AF208224; AAF17100.1; -		
DR	HSSP; P56771.2AAA.		
DR	InterPro: IPR000461; Alpha-amylase.		
DR	InterPro: IPR002044; CBD_4.		
DR	pfam: PF00128; alpha-amylase; 1.		
DR	pfam: PF00686; CBD_4; 1.		
DR	PRINTS; PR00110; ALPHAMYLASE.		
DR	PRODOM; PD001568; CBD_4; 1.		
SO	SEQUENCE 623 AA; 68598 MW; 822F002C37F5A9A4 CRC64;		

Query Match	64.3%	Score 1725.5;	DB 3;	Length 623;
Best Local Similarity	63.7%	Pred. No. 1.5e-109;		
Matches 309;	Conservative 68;	Mismatches 107;	Indels 1;	Gaps 1

[illegible]

```

Db      70 YIQDMGFAIRWITPTEBIPROYTAGVGFHGYMOKNITGVDTNLGTRADIRALSEALHDR 1299
Qy      131 GMYLWMDVYVANHMGYDAGSSVDY SVFKPSSODY FHPEFCIQNYEOTQVEDCWLGDNT 190
      130 GMYLMDVYVANHMSYCGGSGSTFISIFLPEDSASYPHSYCANINNDQMOWENECFLGDT 1899
Qy      191 VSLPRLDTTKDYKKEMEDWQSLYSN SIDGLRIDPYKHQKQFWPGYKACGYTCGE 250
      190 VSLTDLNMQSSSEVRBIWMDIEDIVANSYDGLRIDPYKHHEKDFWPGYIDAAGVSVGE 249
Qy      251 YLDGDPATCPONMMDVLNVPIYPLLNFKKTSQSMODLYMIMTVKSDGCDSTLLG 310
      250 ITHSPATICEPQDMQMDVMMNPPIYPLLNFKKSSQSSMDLYMIMTVANSNCRDPLL 309
Qy      311 TFEVHNHDPREFASTYNDIALKNAAFIILNDGIPRIYAGOEHOYAGCNPANREATWLS 370
      310 NTEIHNHDPREPNNTYPPDMRSKKNLAFLEFLTDGPIYIYAGOEHOYSSNDPNYKREPWMS 3699
Qy      371 GYPDSELYKLIASANAIARNYAISKDQGFVTVYKKNMPYIKDDPTIANKRGKDGSOIVILS 430
      370 SYSTSESLYKFIATNKRIRKLAIKSDSYLSRMTPEYSDSNYIANKKSGGGSVLLTLN 429
Qy      431 NKGASGDSYTTISLSAGYTTAGQOLTEYIGCTTVYVSGDNVPVPMAGCLPRVILPYEKLKA 490
      430 NIGTISIGSYTFEDLYDHGNSGANLVELYTCSSYOVGNSGAISIPMTSGCLPRVILPAAWVS 489
Qy      491 GSKIC 495
      490 GSGIC 494

```

RESULT	3	00117	
ID	Q01117		
AC	Q01117	PRELIMINARY;	PTT: 624 AA.
DT	01-NOV-1998 (TREMBLrel. 08, Created)		
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 11, Last annotation update)		
DE	ALPHA-AMYLASE PRECURSOR (PC 3.2.1.1)		
DE	(1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE) (LKA1).		
GN	LKA1.		
OS	Lipomyces kononenkoae.		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Lipomycetaceae; Lipomyces.		
OX	NCBI_TaxID=34357;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-IGC4052B;		
RC	MEDLINE=96105202; PubMed=8529895;		
RA	Steyn A.J.C., Marum J., Pretorius I.S.;		
RT	"Cloning, sequence analysis and expression in yeasts of a cDNA		
RT	containing a lipomyces kononenkoae alpha-amylase-encoding gene.";		
RL	Gene 166:65-71(1995).		
RN	[2]		
RP	SEQUENCE OF 29-44.		
RC	STRAIN-IGC4052B;		
RC	MEDLINE=96132108; PubMed=8593683;		
RA	Steyn A.J.C., Pretorius I.S.;		
RT	"Characterization of a novel alpha-amylase from lipomyces kononenkoae		
RT	and expression of its gene (LKA1) in Saccharomyces cerevisiae.";		
RL	Curr. Genet. 28:526-533(1995).		
CC	-1- FUNCTION: LIBERATES REDUCING GROUPS FROM POLYMERS CONTAINING BOTH		
CC	1,4-ALPHA AND 1,6-ALPHA BONDS.		
CC	-1- CATALYTIC ACTIVITY: ENDOTHEROLYSIS OF 1,4-ALPHA-GLUCOSIDIC		
CC	LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.		
CC	-1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY		
CC	SIMILARITY).		
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR.		
CC	-1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO		
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY.		
DR	EMBL: U03076; AAC49622.1; ALT_INIT.		
DR	HSSP; P10529; 7TAA.		


```
RESULT 7
ID 092394 PRELIMINARY; PRT; 631 AA.
AC 092394;
DT 01-FEB-1997 (TReMBrel. 02, Created)
DT 01-FEB-1997 (TReMBrel. 02, last sequence update)
DE 01-JUN-2001 (TReMBrel. 17, last annotation update)
DE ALPHA-AMYLASE PRECURSOR.
GN AMY-CS2.
OS Cryptococcus sp. S-2.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
OC Mitosporic Hymenomycetes; Cryptococcus.
OC NCBI_TaxID=87049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S-2;
RX MEDLINE=96431120; PubMed=8836148;
RA Iefuji H., Chino M., Kato M., Iimura Y.;
RT "Raw-starch-digesting and thermostable alpha-amylase from the yeast
RT Cryptococcus sp. S-2: purification, characterization, cloning and
RT sequencing.";
RL Biochem. J. 318:989-996(1996).
DR EMBL; D83541; BAA12011.1; -.
DR EMBL; D83540; BAA12010.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro: IPR000461; Alpha_amylase.
DR InterPro: IPR002044; CBD_4.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF00686; CBD_4; 1.
DR ProDom; PD001568; CBD_4; 1.
DR Signal.
KW SIGNAL.
FT CHAIN 1 20 ALPHA-AMYLASE.
FT SIGNAL 21 631
SQ SEQUENCE 631 AA; 67658 MW; 8196B7B6E1D707E5 CRC64;

Query Match 47.1%; Score 1264; DB 3; Length 631;
Best Local Similarity 49.2%; Pred. No. 4.4e-78;
Matches 251; Conservative 84; Mismatches 153; Indels 22; Gaps 9;

OY 9 LYGQVAPALAA--TRADRSOSITFLIDRRARIDGSTATACNADQ--KTCGSTNG 64
DB 7 LAGALLASLGVLVAGLSPEWRSOSITGVYDRALDNGMSPSCQSSELTINICNGTFAG 66
OY 65 IIDKLDYIQMGFTAIWTPVTAQLP-----QTAYGDAHYGWODIYSLNENGTAD 119
DB 67 IIDKLDYIQMGFTAIWTPVTAQLP-----QTAYGDAHYGWODIYSLNENGTAD 126
OY 120 LKAISALHERGATLMDVYVANHNGY-----DGAGSSVDYSEVFKPFSSQDYFHPFC 172
DB 127 LTDSNASHSRGMYLMDVYVANHNGY-----DGAGSSVDYSEVFKPFSSQDYFHPFC 186
OY 173 QNVEDOTQVECCWLGDMVSLPDDDTTKDYVKNEMWDWGSIVSNSTIDLRIDYKHQY 232
DB 187 -DYNNRRTSIDLWEGDELVELYLDRTEDSDVQSIFNSMINLQITNIDSLRIDSLQDS 245
OY 233 KDFEYGVKAA-GVYCIGEVLGDPAVTCYQVY-MDGVLMYPIYYPPLNAFKSTGSM 290
DB 246 SEFFEGFQAAGMYMGEVFNQSPSYCYQAGMGVLMYPMFEYITNAFOTSSGMS 305
OY 291 DLYNNINNVKSCDPSDTLIGFEVENHNNPRPASTNDIALAKVNAFTINDGIPITV 350
DB 306 QLAGISMSQSDSTTLIGFLENQDNPREPSQSLTRAQNALIAQDSIPITVYG 365
OY 351 QEOHYAGNDPANREATWLS-GYPTDSELYKLASANAIRNVAISKDTGVTAKNPIYK 409
DB 366 QEOHLSGSGVPLNREALMTSGDYDTSSPLXEMITTVQRLTAKIQNGGVTVYKIVPT 425
OY 410 DDTITAMKKGIDGSOYITILSNKASDSTYLSLGA--GYTAGQDLTEVIGCTTV 467
DB 426 DSNHIVTRKNGSGYQIVGVYTNVSGSSSTLSISSETGFQASEPYMDVLSCTLYHTGT 485
OY 468 DGNVPVPMAGGLPRVLYPTEKLAGSKICSS 497
```

```
DB 486 DGSLSFTMTGGLPRVFNATATALESSICTT 515

RESULT 8
ID 074922 PRELIMINARY; PRT; 625 AA.
AC 074922;
DT 01-NOV-1998 (TReMBrel. 08, Created)
DT 01-NOV-1998 (TReMBrel. 08, last sequence update)
DE 01-DEC-2001 (TReMBrel. 19, last annotation update)
DE ALPHA-AMYLASE.
GN SPEC757.12.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OC NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Lyne M., Rajandream M.A., Barrell B.G., Bothe G., Pohl T;
RL submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031825; CA21237.1; -.
DR HSSP; P10529; 7TAA.
DR InterPro: IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
SQ SEQUENCE 625 AA; 68087 MW; 30A58CE6DB0C07B3 CRC64;

Query Match 35.9%; Score 963.5; DB 3; Length 625;
Best Local Similarity 40.7%; Pred. No. 1.3e-57;
Matches 201; Conservative 83; Mismatches 195; Indels 15; Gaps 9;

OY 2 VAMNSLF-LYGQVAPALAA--TRADRSOSITFLIDRRARIDGSTATACNADQKTCG 60
DB 6 IALPSLALFGLP---SLAKSSEWDRITLYQVYDRFA-VSDMTPPCSPDSSYCG 61
OY 61 TWOGIIRKLDYIQMGFTAIWTPVTAQLPOT--TAYGDAHYGWODIYSLNENGTAD 119
DB 62 TWSGIRKLDYIQMGFTAIWTPVTAQLPOT--TAYGDAHYGWODIYSLNENGTAD 121
OY 120 LKAISALHERGATLMDVYVANHNGYDGAGSSVDYSEVFKPFSSQDYFHPFCIQ-NYED 178
DB 122 LIDLTDMHNDMMIMFDALANSWAIIPGPDNITSNLYVFNDSSYFHHYCMWDIGSN 181
OY 179 TQVEDCWTGDMTVSLPDDDTTKDYVKNEMWDWGSIVSNSTIDLRIDYKHQKDFP 238
DB 182 TQVEDCWTGDMTVSLPDDDTTKDYVKNEMWDWGSIVSNSTIDLRIDYKHQKDFP 241
OY 239 YNKAAGVYCIGEVLGDPAVTCYQVYVLMYPIYYPPLNAFKSTGSMIDLYNMINT 298
DB 242 YTSAAAGVYCIGEVLGDPAVTCYQVYVLMYPIYYPPLNAFKSTGSMIDLYNMINT 301
OY 299 VKSDC--PDSITLLTFEVENHNNPRPASTNDIALAKVNAFTINDGIPITVAG 356
DB 302 FQACCEGQDMVYVGNFLENHDLPRYTSITNDTSQDGAIFLLHNGGIPLIYGEQ 361
OY 357 GGNP-PANREATWLSGPTDSELYKLASANAIRNVAISKDTGVTAKNPIYKMDPT 415
DB 362 GGSOTPENRALMNYGDTIDTANYYQITRTIALRLKQALISDSBWTDSYLDIRAH 421
OY 416 MRKGTDSQIVYITILSNKASDSTYLSLGA--GYTAGQDLTEVIGCTTVVGS 475
DB 422 VRKG---DVLGVYTNVSGSSSTLSISSETGFQASEPYMDVLSCTLYHTGT 476
OY 476 AGGLPRVLYPTEKLAGSKICSS 489
DB 477 VSGLPVYTPPEASL 490

RESULT 9
ID 013996
```

```

ID 013996 PRELIMINARY; PRT; 491 AA.
AC 013996;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DE 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE PUTATIVE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1).
GN SPAC27E2.01.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -! CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -! SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR EMBL; 298978; CAB11675.1; -.
DR HSSP; P10529; 77RA.
DR InterPro; IPR000461; Alpha-amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
KW Glycoprotein; Signal.
KM SIGNAL 1 491 POTENTIAL.
FT CHAIN ? 491 BY SIMILARITY.
FT ACT_SITE 221 221 BY SIMILARITY.
FT ACT_SITE 225 225 BY SIMILARITY.
FT ACT_SITE 313 313 BY SIMILARITY.
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 491 AA; 56678 MW; 2013808DC0EB581D CRC64;

```

```

Query Match 34.3%; Score 920.5; DB 3; Length 491;
Best Local Similarity 38.7%; Pred. No. 8.2e-55;
Matches 181; Conservative 79; Mismatches 185; Indels 23; Gaps 6;

```

```

OY 26 WKSOSIYFLIDRFRKGTGTTATCNTADQKYCGGTGMOGIIIDKDYIOGSGFTAIWTPY 85
DB 24 WKRQCTIYQILDHFDHCHTAP--STGRMYLGITWKGIIQKDYIOGSGCTAWISPI 81
OY 86 TADPOTTAAGDAGHYGWOQDIYSLNENYGTADDLKALSSALHBRGMYLWVYVANHGY 145
DB 82 VKNIGVGYGEAYHGYWAEDLTOLNPHFTKODLTETLVDLHRRNMCMIDIYVNHMAH 141
OY 146 DGAGSVYVYKPFSSODYPHPCFIQNYEDQOVEDCWMGDMFTVSLPDLDTKDYVN 205
DB 142 AG-DSFIDISKYAPNNSSHHPRKFLNHYDIDWDCETAMIGDVSLSMDRTEDQEVHN 200
OY 206 EYDWVGLSVNYSIDGLRIDTVKHVKORDFWPGYNKAGVYCIGEVLDGDPAYTCPYQNV 265
DB 201 FQONMIRLDIQYHFDGRIIDAKHYKEFPFPIAANVYAFGEVYHGPCKFIKYLEY 260
OY 266 MGVGLNPIYPLLAFASTSGSMDDL-NKINIVYKSCPPSTLSLTFVEVHNDPRASY 324
DB 261 IESAANPYLYQIENFTFPPKQSMNIFOKALILEARASTMOTYTLIGFTEHVDHDFLNR 320
OY 325 TNDIALAKVAAFIILNDGIPRIYAGQEQHAYAGNDPANRATWLSGYPDSELYKLIAS 384
DB 321 STDYSLICNTLTLFLFTGIGIPRIIFGQEQOMYAGGIDPENRALMTSNVNOGNPIFOPLK 380
OY 385 ANAIRNVAISKDTGVTYKKNPDIYDDTTIAMRKGTDSQIVTLTSLNKGASGDSYTL--- 441
DB 381 LTKRFELVDNYSGETTESNMLFVNEHYVYFR---PGVILVVSNAAGSSVDVISAEE 436
OY 442 -----SLSGAGYTAGOOLTEVIGCTIVTGSDSGVNVPYPMAGGLPVL 483
DB 437 SITTESLEFDIVLSGSDF-----SLPTEDSSTISNMLEFSPRYL 478

```

```

RESULT 10
OY7S9
ID 09Y7S9 PRELIMINARY; PRT; 564 AA.
AC 09Y7S9;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE PROBABLE ALPHA-AMYLASE C63.02C PRECURSOR (EC 3.2.1.1) (1.4-ALPHA-D-
DE GLUCAN GLUCANAMYLASE).
GN SPC63.02C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Murphy L., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -! CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -! COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
CC SIMILARITY).
CC -! SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR EMBL; AL049522; CAB4006.1; -.
DR HSSP; P10529; 77RA.
DR InterPro; IPR000461; Alpha-amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
KW Calcium; Glycoprotein; Signal.
KM SIGNAL 1 21 POTENTIAL.
FT CHAIN 1 21 PROBABLE ALPHA-AMYLASE C63.02C.
FT DOMAIN 537 540 POLY-SER.
FT ACT_SITE 229 229 BY SIMILARITY.
FT ACT_SITE 322 322 BY SIMILARITY.
FT DISULFID 51 59 BY SIMILARITY.
FT DISULFID 172 188 BY SIMILARITY.
FT DISULFID 263 306 BY SIMILARITY.
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 564 AA; 63205 MW; 1229PDAAE702FDD0 CRC64;

```

```

Query Match 33.5%; Score 898.5; DB 3; Length 564;
Best Local Similarity 38.8%; Pred. No. 3.1e-53;
Matches 188; Conservative 85; Mismatches 190; Indels 21; Gaps 7;

```

```

OY 18 ALAAPADWRSOSIYFLIDRFRKGTGTTATCNTADQKYCGGTGMOGIIIDKDYIOGSGF 77
DB 19 ANAGSNAMERKRIIYQIILTRFAVDDSDTNPDCPDANOTCGGFWKCIENKDIYEDMGF 78
OY 78 TALWITPYTAOLP-QTTAYGDAYHYGWOQDIYSLNENYGTADDLKALSSALHBRGMYLW 136
DB 79 NAIWISPIDKNIEGIDGAGYAHGYWNTDYESINHEHFGTEEDLVSLITAAHAGIWMVL 138
OY 137 DVYVANHMGYDAGSSVDVSVKPFSSODYPHPCFI--QNYVEDQOVEDCWMGLDNTVSLP 194
DB 139 DSVSMALAPPLAADSSLNPFNKESYFHPYOLIMDITIDETNVMDCW-ODSGVLLA 197
OY 195 DLDITKDYVKNKEMWDWVGLSVNYSIDGLRIDTVKHVKORDFWPGYNKAGVYCIGEVLDG 254
DB 199 DLDVSSSVSYSLSDHFSKLSISKYDPOGLRIDVAKMNNYFFEPFVDATGVYSVGEVFSY 257
OY 255 DPAYTCPIYONWGVGLNPIYPLLAFAKSTSGSMDDLNNKINIVKSDCP--DSTLIGTF 312

```



```

Db 258 DEDYMCSTYSVLPQVNTFLQLYNFSFATGACFTLLIPYQEVNVAASCKYDSTLMLTF 317
OY 313 VENHNDPRFASSTNDIALAKNVAFAFLINDGIPITVAGQEOHYAGNPNAREATWLSGY 372
Db 318 IENHDLRPPTYSQSOQMGALSTVLLMDGIPSTFYGOEGNGEDPANPALMLTDY 377
OY 373 PPDSELYKLIASANAIRNVAISKDGFVYTKNMPYIKDDTTIARRKGTND---GSOIYT 427
Db 378 DQSNPYTVYIKTMVAFRKFTVITQDDPWT-----STQGISASVDHYGKNDVLY 428
OY 428 ILSKAGSDSTLSLSGAGTACGOLTEVIGCTVTVGSDGNVPNAGGLPRVLYPTE 487
Db 429 MNNMGVTV-NMLTIYEVTNTYANNEVSDVFGHRTLVGADKXTLTASMTNGYPLIMYDHS 487
OY 488 KLAG 491
Db 488 KMSG 491

```

```

RESULT 11
060051
ID 060051 PRELIMINARY; PRT; 482 AA.
AC 060051.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1).
GN AMYV.
OS Thermocotilomyces vulgaris.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Thermocotilomyces.
OX NCBI_TaxID=2026;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K94.
RA MEDLINE=95031040; PubMed=7944369;
RA Hotemeister B., Koenig S., Hoang V., Engel J., Mayer G., Hansen G.,
RA Hotemeister J.;
RA "The gene amyE(TV1) codes for a nonquicogenic alpha-amylase from
RT Thermocotilomyces vulgaris 94-2A in Bacillus subtilis."
RL Appl. Environ. Microbiol. 60:3381-3389(1994).
DR EMBL: X69807; CAI49465.1; -.
DR HSSP: P10529; 7TAA.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
DR Pfam: PF02806; alpha-amylase-C; 1.
KM Signal: Hydrolase; Glycosidase.
FT SIGNAL 1 29
FT CHAIN 30 482 POTENTIAL.
FT SEQUENCE 482 AA; 54483 MW; D8F98C370B579025 CRC64;

```

Query Match 27.5%; Score 737; DB 2; Length 482;
 Best Local Similarity 36.8%; Pred. No. 2.6e-42;
 Matches 185; Conservative 77; Mismatches 189; Indels 52; Gaps 15;

```

OY 1 MYAMWSLEFLYGQVAAAPALAAATPADRMSOSIYFLTDFRFRDGTST---ATCMTAD-Q 55
Db 10 LVCLLSAVLMFRSTAPRTVALSPADMOEKSTIFTITDFDNGDSNDNYGFFNADKSDPR 69
OY 56 KYCGGTWQGIIDKLDYIOGMGFTAIWTPVTAQPLPQTTAYGDAYHGYQODIYSLNENYG 115
Db 70 KKHGDFQGIINRLDYIKMGFTAIWTPVMO-----KSANAYHGYWTYDFYSVDHLG 124
OY 116 TADDLKALLSAHHERGMLNDVYVANHNGYDAGASSVYVSEKFPSSQDYHHPFCFIONY 175
Db 125 SNAKIQEIVRAHNAKGISVMDLVVANHNGT-DFQPSSEFKA--FPDKADWYHNGYDQDM 180
OY 176 EDQTOVEDCMGLDNTVSLPDLDTTKDYVKNEMVYDGLSLNYSIDGLRITVHYVQDF 235
Db 181 NQGMWVEN---GD-IAIGDLDLNDQNPRAVATELKNMIANLVTGTGVDGLRVTVVAVHVKWF 236
OY 236 WFGYKKAAGVYICIGEVLDGDGPATYTCYQNVMDGVLYNPDIYPLNLF--KSTSGSMIDL 293

```

```

Db 237 WREDFGAANFTLTGEVPHGCDPAVYGVDTYNLDVNDLDEPMYTTIKNFGQDQSRITADRX 296
OY 294 NMATVNSDCPDSTLLTGFVENHNDPRFASSTNDIALA-----KNVAAPITLLDGI 345
Db 297 AODRYRYKNPMTN---GLFIDNHVPREFLNDSKPGASMDKWPOLKAALGFMTTGPID 352
OY 346 IIVAGQEOHYAGNPNAREATWLSGYPTDSELYKLIASANAIRNVAISKDGFVYTKN 405
Db 353 ILYGTRQGSFGADPYNRD---MWFNDKHELYTAKLNTYRNHVALRKSTQAEK-- 407
OY 406 PIYKDDTTIARRKGTNDGSOIYTLISNKGASDSTLSLSGAGYTA-GQOLTEVIGCTVY 464
Db 408 --WVDDTFYAFQBSYNGDEVVYMIN--SMNSQTRVPVNDIVANGQILYNNMGIDKYT 462
OY 465 VGSQGNVPNAGGLPRVLYPTE 487
Db 463 VNN-----GSITATLTGPKE 476

```

```

RESULT 12
042918
ID 042918 PRELIMINARY; PRT; 774 AA.
AC 042918;
DT 01-JAN-1999 (TREMBLrel. 09, Created)
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE ALPHA-AMYLASE C16A3.13 PRECURSOR (EC 3.2.1.1) (1.4-ALPHA-D-
DE GLUCAN GLUCANOHYDROLASE).
GN SPBC16A3.13.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Rajandream M.A., Barrell B.G., Beck A., Reinhardt R.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOLYOLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
CC SIMILARITY).
CC -1- SIMILARITY: TO THE FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS
CC THE ALPHA-AMYLASE FAMILY.
DR EMBL: AL021748; CAI16864.1; -.
DR HSSP: P56271; 2AAA.
DR InterPro: IPR000461; Alpha-amylase.
DR Pfam: PF00128; alpha-amylase; 1.
KM Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
KM Calcium; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 774
FT ACT SITE 461 461 PUTATIVE ALPHA-AMYLASE C16A3.13.
FT ACT SITE 553 553 BY SIMILARITY.
FT DISULFID 52 60 BY SIMILARITY.
FT DISULFID 229 314 BY SIMILARITY.
FT DISULFID 495 539 BY SIMILARITY.
FT DISULFID 690 724 BY SIMILARITY.
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 408 408 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 774 AA; 89459 MW; 92CEAAE45664269B CRC64;

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:48:50 ; Search time 158.68 Seconds
(without alignments)
348.593 Million cell updates/sec

Title: US-09-710-339-2

Perfect score: 2684
Sequence: 1 MVAMWSLFYGLGVADPALA.....LPRLVPTKLAGSKICSSS 498

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802:*
1: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2684	100.0	498	22	AAB84206	Amino acid sequenc
2	2572	95.8	478	16	AA872450	Aspergillus oryzae
3	2572	95.8	478	16	AA878270	Aspergillus oryzae
4	2572	95.8	478	17	AA814500	Aspergillus oryzae
5	2552	95.1	478	15	AA846055	Aspergillus oryzae
6	2503.5	93.3	478	16	AA878025	Mutant alpha-amyl
7	1666.5	62.1	493	17	AA888212	Mature taka-amyl
8	1449	54.0	512	11	AA807574	Alpha-amylase. Tr
9	1400.5	52.2	511	9	AA81180	Alpha-amylase enco
10	1398.5	52.1	511	9	AA81161	Sequence of alpha-
11	1397	52.0	468	13	AA824136	Recombinant alpha- Alpha-amylase vari

12	1396.5	52.0	494	8	AA870571	Alpha-amylase gene
13	1392	51.9	468	15	AA863184	Variant alpha amyl
14	1391	51.8	468	15	AA863185	Variant alpha amyl
15	1387	51.7	468	15	AA863187	Variant alpha amyl
16	1338	49.9	452	15	AA863186	Variant alpha amyl
17	453	16.9	719	22	AA874220	Bacillus maltogeni
18	452.5	16.9	712	12	AA810051	Cyclomaltoextrin
19	452	16.8	719	22	AA874219	Bacillus maltogeni
20	450	16.8	719	20	AA830621	Amino acid sequenc
21	450	16.8	719	20	AA831731	Bacillus sp. malo
22	450	16.8	719	21	AA894271	Bacillus maltogeni
23	450	16.8	719	22	AA859556	Bacillus maltogeni
24	450	16.8	719	22	AA874216	Bacillus maltogeni
25	449.5	16.7	656	11	AA806109	Sequence of cyclom
26	446	16.6	719	22	AA874217	Bacillus maltogeni
27	446	16.6	719	22	AA874218	Bacillus maltogeni
28	446	16.6	719	22	AA874222	Bacillus maltogeni
29	446	16.6	719	22	AA874223	Sequence of cyclom
30	444.5	16.6	713	11	AA806110	Cyclomaltoextrin
31	444.5	16.6	713	12	AA810052	Bacillus maltogeni
32	444	16.5	719	22	AA874221	Bacillus maltogeni
33	432.5	16.1	683	17	AA817636	Thermoanaerobacter
34	432	16.1	682	17	AA817637	Thermoanaerobacter
35	430.5	16.0	683	17	AA817640	Thermoanaerobacter
36	428.5	16.0	655	13	AA820139	Sequence encoded b
37	427.5	15.9	681	17	AA817615	Thermoanaerobacter
38	427.5	15.9	683	17	AA817641	Thermoanaerobacter
39	427	15.9	526	19	AA844845	Cyclomaltoextrin
40	426.5	15.9	683	17	AA806772	Wild type cyclomal
41	426.5	15.9	683	17	AA817646	CGTase variant 146
42	426.5	15.9	683	17	AA817647	CGTase variant 145
43	426.5	15.9	683	17	AA817648	CGTase variant 145
44	426.5	15.9	683	17	AA817651	CGTase variant 145
45	426.5	15.9	683	17	AA817622	Thermoanaerobacter

ALIGNMENTS

RESULT 1
ID AAB84206 standard; Protein: 498 AA.
XX AAB84206;
AC
XX
DF 06-AUG-2001 (first entry)
XX
DE Amino acid sequence of a fungamyl-like alpha-amylase.
XX
KW Fungamyl-like alpha-amylase; glucoamylase; dextrinifascation; maltose;
KW alcohol; starch; dough improver; brewing; starch liquification.
XX
OS Aspergillus oryzae.
XX
XX
PN WO200134784-A1.
XX
PD 17-MAY-2001.
XX
PE 10-NOV-2000; 2000WO-DK00626.
XX
PR 10-NOV-1999; 99DK-0001617.
XX
PA (NOVO) NOVOZYMES AS.
XX
PI Bisgard-Frantzen H, Svendsen A, Pedersen S;
XX
DR WPI: 2001-367478/38.
XX
N-PSDB: AAF90208.
XX
PT New variant of Fungamyl-like alpha-amylase, useful for production of
PT maltose syrups, includes mutations that improve stability against heat
and acidic pH -
XX

PS Claim 1: Page 42-45; 49pp; English.
XX
CC The present sequence represents a fungamyl-like alpha-amylase. The
CC specification describes variants of this fungamyl-like alpha-amylase,
CC which have an alteration in one the amino acid regions 98-110, 150-160,
CC 161-167, 280-288, 448-455, and 468-475. Each alteration is a deletion
CC or substitution of an amino acid or an insertion of an amino acid
CC downstream of a particular position. The variants retain alpha-amylase
CC activity, and have better heat stability and/or stability at acidic pH,
CC relative to wild-type enzyme. The variants can therefore be used at
CC higher temperatures (more efficient conversion or faster reaction, and
CC have reduced need for cooling and reduced risk of contamination). The
CC variants may also be used in conjunction with other enzymes,
CC particularly glucoamylase during dextrinisation. The variants are
CC used to produce syrups, particularly of high maltose content, or alcohol,
CC from starch; as dough improver for baked goods; in brewing, to increase
CC fermentability of the wort; and for liquefaction of starch.
XX
SQ Sequence 498 AA:

Query Match 100.0%; Score 2684; DB 22; Length 498;
Best Local Similarity 100.0%; Pred. No. 8.3e-227;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MWAMWSLFLYGLQVAPALATPADMRSSQSYFLTDREARTDGGSTATCMTADOKYCGG 60
DB 1 mwamwsflflyqlvaapalaapadmrsgsyfltdrtatdgsstctctadqkkygg 60
OY 61 TWOGIIDLKIDYIQMGFTAIWITPVAOLPQTAYGDAYHGYWQDIYSLNENGTADDL 120
DB 61 twogiidlkidyiqmgftaiwitpvaolpqtaygdayhgywqdiyslnenytaddl 120
OY 61 twgjidkidyiqmgftaiwitpvaolpqtaygdayhgywqdiyslnenytaddl 120
DB 61 twgjidkidyiqmgftaiwitpvaolpqtaygdayhgywqdiyslnenytaddl 120
OY 121 KALSSALHERGMYLWADVANAHEMGDAGSSVDYVFEKPESSODYFHPFCFIONEDOTO 180
DB 121 kalsalhergmylmwadvanahemgdagsvdyvfkpfssqdyfhpfcfionyedq 180
OY 121 kalsalhergmylmwadvanahemgdagsvdyvfkpfssqdyfhpfcfionyedq 180
DB 121 kalsalhergmylmwadvanahemgdagsvdyvfkpfssqdyfhpfcfionyedq 180
OY 181 VEDCWLGDNTVSLPDLDTTKDYVKNEMVDWVSLVSNTSIDGLRIDTYKHVKDFWPEYN 240
DB 181 vedcwlgdntvslpdltdtkdyvknemv dwvslvsnsidglridtykhvkdfwpeyn 240
OY 181 vedcwlgdntvslpdltdtkdyvknemv dwvslvsnsidglridtykhvkdfwpeyn 240
DB 181 vedcwlgdntvslpdltdtkdyvknemv dwvslvsnsidglridtykhvkdfwpeyn 240
OY 241 KAAGVYCIQEVLDGDPATYCPYQNWGVLYNPPIYPLINAFKSTSGSKMDLYNMINTVK 300
DB 241 kaagvyciqevldgdpaycipyqnmvgvlynpipyplinafkstsgsmddlynmintvk 300
OY 241 kaagvyciqevldgdpaycipyqnmvgvlynpipyplinafkstsgsmddlynmintvk 300
DB 241 kaagvyciqevldgdpaycipyqnmvgvlynpipyplinafkstsgsmddlynmintvk 300
OY 301 SOCPDSTLTGTEVENHNDNRFASYNDDIALAKNVAFIINDGIPITAYAGOEQHYAGND 360
DB 301 scdpdstltgtvenhndnrfasynddialaknvaafiiindgipilyaqeqhyagnd 360
OY 301 scdpdstltgtvenhndnrfasynddialaknvaafiiindgipilyaqeqhyagnd 360
DB 301 scdpdstltgtvenhndnrfasynddialaknvaafiiindgipilyaqeqhyagnd 360
OY 361 PANREATWLSGYPTDSELEKLIASANAIRNVAISKDTGCVTKKMPYKDDTTIARRKGT 420
DB 361 panreatwlsyptdseleykliasanairnvaistdgtcvtkmnpykddttiamrkt 420
OY 361 panreatwlsyptdseleykliasanairnvaistdgtcvtkmnpykddttiamrkt 420
DB 361 panreatwlsyptdseleykliasanairnvaistdgtcvtkmnpykddttiamrkt 420
OY 421 DGSQIVTILSNKAGSDSYTSLSGAGYTAGOQLTEVIGCTTVYVGSNGNVPYPAAGLP 480
DB 421 dgsqivtillsnkagsdysytlisgagyttagoqltevigcttvvyvgsdgnvypaaglp 480
OY 421 dgsqivtillsnkagsdysytlisgagyttagoqltevigcttvvyvgsdgnvypaaglp 480
DB 421 dgsqivtillsnkagsdysytlisgagyttagoqltevigcttvvyvgsdgnvypaaglp 480
OY 481 RVLYPEXLAGSKICSS 498
DB 481 rvlypeklagskicss 498

RESULT 2

AA72450
ID AAR72450 standard; Protein: 478 AA.

XX
AC AAR72450;

DT 01-DEC-1995 (first entry)

DE Aspergillus oryzae alpha amylase (mature protein).

XX Alpha amylase; variant; enzyme; detergent; additive; dishwashing;

KW washing; Bacillus licheniformis; Bacillus amyloliquefaciens;
KW Bacillus stearothermophilus; dyeing; bleaching; scouring; textile;
KW thermostable.

XX Aspergillus oryzae.

OS WO9510603-A.

PN 20-APR-1995.

PD 05-OCT-1994; 94WO-DK00370.

PF 08-OCT-1993; 93DK-0001133.

PR 02-FEB-1994; 94DK-0000140.

XX (NOVO) NOVO-NORDISK AS.

XX Svendsen A, Thellersen M;

PI Van der zee P, Bisgardfrantzen H, Borchert T;

XX WPI; 1995-161790/21.

PT New Bacillus derived alpha-amylase variants - having amino acid
PT modifications to improve washing and/or dishwashing performance

XX Disclosure; Page 75-76; 105pp; English.

XX Variant alpha amylase enzymes which have improved washing and/or
XX as detergent additives. The enzymes have one or more amino acid
XX residues added, deleted or substituted. The variants can also be
XX used for textile desizing prior to scouring, bleaching and dyeing.

CC The variants have improved thermostability, acid/alkaline stability,
CC low temperature optimum; pH optimum; higher hydrolysis velocity and
CC improved tolerance to other composition constituents, e.g. oxidation
CC agents.

XX Sequence 478 AA:

Query Match 95.8%; Score 2572; DB 16; Length 478;
Best Local Similarity 99.8%; Pred. No. 5.3e-217;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 21 ATPADMRSSQSYFLTDREARTDGGSTATCMTADOKYCGGTWOGIIDLKIDYIQMGFTAI 80
DB 1 atpdmrsgsyfltdrtatdgsstctctadqkkygwtwogiidlkidyiqmgftai 80
OY 81 WITPVAOLPQTAYGDAYHGYWQDIYSLNENGTADDLKALSSALHERGMYLWADVANA 140
DB 61 witpvaolpqtaygdayhgywqdiyslnenytaddlkalsalhergmylmwadvana 120
OY 61 witpvaolpqtaygdayhgywqdiyslnenytaddlkalsalhergmylmwadvana 120
DB 61 witpvaolpqtaygdayhgywqdiyslnenytaddlkalsalhergmylmwadvana 120
OY 141 NHMGYDGGSSVDYVFEKPESSODYFHPFCFIONEDOTVEDCWLGDNTVSLPDLDTTK 200
DB 141 nhmgydgssvdyvfkpfssqdyfhpfcfionyedqvedcwlgdntvslpdltdtk 180
OY 141 nhmgydgssvdyvfkpfssqdyfhpfcfionyedqvedcwlgdntvslpdltdtk 180
DB 141 nhmgydgssvdyvfkpfssqdyfhpfcfionyedqvedcwlgdntvslpdltdtk 180
OY 201 DVVKNEMVDWVSLVSNTSIDGLRIDTYKHVKDFWPEYNKAAGVYCIQEVLDGDPATYC 260
DB 181 dvvknemv dwvslvsnsidglridtykhvkdfwpeynkaagvyciqevldgdpaytc 240
OY 201 dvvknemv dwvslvsnsidglridtykhvkdfwpeynkaagvyciqevldgdpaytc 240
DB 201 dvvknemv dwvslvsnsidglridtykhvkdfwpeynkaagvyciqevldgdpaytc 240
OY 261 PYQNWGVLYNPPIYPLINAFKSTSGSKMDLYNMINTVKSDCPDSTLTGTEVENHNDN 320
DB 241 pyqnmvgvlynpipyplinafkstsgsmddlynmintvksdcpdstltgtvenhndn 300
OY 241 pyqnmvgvlynpipyplinafkstsgsmddlynmintvksdcpdstltgtvenhndn 300
DB 241 pyqnmvgvlynpipyplinafkstsgsmddlynmintvksdcpdstltgtvenhndn 300
OY 321 FASTYNDIALAKNVAFIINDGIPITAYAGOEQHYAGNDPANREATWLSGYPTDSELYK 380
DB 301 fastyndialaknvaafiiindgipilyaqeqhyagndpanreatwlsyptdseleyk 360
OY 381 LIASANAIRNVAISKDTGCVTKKMPYKDDTTIARRKGTGDSQIVTILSNKAGSDSYT 440
DB 361 liasanairnvaistdgtcvtkmnpykddttiamrktgdsqivtillsnkagsdysyl 420
OY 441 LSLSGAGYTAGOQLTEVIGCTTVYVGSNGNVPYPAAGLP RVLYPEXLAGSKICSS 498
DB 441 lslsgagyttagoqltevigcttvvyvgsdgnvypaaglp rvlypeklagskicss 498

Db 421 lsisgagytgqqltevigctvtvsgdgnvpvmagjprvlypteklagskicss 478

RESULT 3

AAR78270

ID AAR78270 standard; Protein; 478 AA.

XX AAR78270;

DT 17-JAN-1996 (first entry)

DE Aspergillus oryzae alpha amylase (mature protein).

KW Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;

KW starch; thermostable; methionine; Bacillus licheniformis;

KW Bacillus amyloliquefaciens; Bacillus stearothermophilus;

KW Aspergillus oryzae.

OS Aspergillus oryzae.

PN W09521247-A1.

PD 10-AUG-1995.

PF 05-OCT-1994; 94MO-DK00371.

PR 02-FEB-1994; 94DK-0000141.

PA (NOVO) NOVO-NORDISK AS.

PI Marcher D, Nilsson TE, Pedersen HH, Toft AH;

DR WPI; 1995-283767/37.

PT Use of an oxidation stable alpha-amylase - for simultaneous desizing and bleaching or scouring of fabrics contg. starch or starch derivs.

PS Disclosure: Page 25-26; 37pp; English.

CC Oxidation stable alpha amylases can be used for the simultaneous desizing and bleaching or scouring of a fabric comprising starch or starch derivatives. They exhibit a better heat stability, especially in the presence of oxidising agents. They are obtained from a parent alpha amylase by replacing one or more methionine residues with any amino acid different from Cys or Met, preferably Leu, Thr, Ala, Gly, Ser, Ile or Asp. The parent alpha amylase is CC pref. derived from a Bacillus species, although alpha amylases of CC fungal origin can also be used. This sequence is the wild type CC (unmodified) alpha amylase of Aspergillus oryzae.

SO Sequence 478 AA;

Query Match 95.8%; Score 2572; DB 16; Length 478;

Best Local Similarity 99.8%; Pred. No. 5.3e-217;

Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPAWRBSQSYIFLLTDFPRTDGTATCNTADQKYGCGTWGQIIXDLXYTQMGFTAI 80
 Db 1 atpadwrsqsyiflltdfartdgsdtatcntadqkycgctwgqikldyiqgmgtfai 60
 QY 81 WTPPTAOLPOTTAGDYGHWGWOODIYSLNBNYGTADLKAALSALHERGMYLWDVYA 140
 Db 61 wtpptaolpotttagdygdhgywgqdlyslnehygtadlkaalsahergmylmvdyva 120
 QY 141 NHMGVDGAGSSVDYSVFEPFSSODYFHPFCFTIONVEDQTOVEDCWLGNPTVSLPDLDTTK 200
 Db 121 nhmgvdgagssvdyvfkpfassqdyfhpfcftionvedqtwedcwlgnptvslpdltdtk 180
 QY 201 DVMKMEWDWNGSLVSNISDGLRDIYKHKQKDFWPEYKNAAGVYICGEVLDDPAVTC 260
 Db 181 dvmknewdwngslvsnysidglridtkvhvkdftpynkaagvycigevldgdpaytc 240

QY 261 PYQNMWDGLVNIPIYPLINAEKSTSGSMDDLYNNINIVKSDCDSTLLGTFVENHNDNR 320
 Db 241 pyqnmwdgvlvniypyplinaekstsgsmddlynninivkscdcpstllgtfvenhndnr 300
 QY 321 FASYNNDIALAKNVAFAFIILNDGIPITIIAGOEOHVAGSNDPANEATLISGYPDSLEYK 380
 Db 301 fasyndndialaknvaefiilndgipilyaggeqiyaggnpanreaklsgypdselk 360
 QY 381 LIASANAIRNVAISKDGTGFVYKKMPYKDDTTIAMRKGTDSQIVTILSNKGAAGSDSYT 440
 Db 361 liasanairnvaiskdgtgfvykkmplykddtliamrkgtsgsqivtllsnkgasgsdyt 420
 QY 441 LLSGAGYTAGGOLTEVIGCTVTWVSGDNVPVMAGLPRVLYPTTEKLASKICSS 498
 Db 421 lsisgagytgqqltevigctvtvsgdgnvpvmagjprvlypteklagskicss 478

RESULT 4

AAW14500

ID AAW14500 standard; Protein; 478 AA.

AC AAW14500;

DT 04-JUN-1997 (first entry)

DE Aspergillus oryzae alpha-amylase (mature protein).

KW alpha-amylase; mature; Bacillus licheniformis; termamyl; fungamyl;

KW Aspergillus oryzae; Bacillus amyloliquefaciens; altered property;

KW calcium dependency; substrate binding; stability; pH optimum;

KW thermostability; cleavage; oligosaccharide substrate; dishwashing;

KW washing; detergent additive; fabric desizing; starch liquefaction;

KW sweetener; ethanol production; variant.

OS Aspergillus oryzae.

XX

Key Location/Qualifiers

FT MISC-difference 13..45

FT /label="loop 1 modification region

FT /note="at least one amino acid residue of a parent

FT alpha-amylase (used as a template for a variant)

FT corresponding to 7-23 of AAW14499 is deleted or

FT replaced with a fragment corresponding to

FT this fragment; claim 33"

FT MISC-difference 14..40

FT /label="loop 1 modification region

FT /note="preferred region where at least one amino acid

FT residue of a parent alpha-amylase (used as a

FT template for a variant) corresponding to 8-18

FT of AAW14499 is deleted or replaced with a

FT fragment corresponding to this fragment; claim

FT 35" MISC-difference 28..42

FT /label="loop 1 modification region

FT /note="at least one amino acid residue of a parent

FT alpha-amylase (used as a template for a variant)

FT corresponding to 12-19 of AAW14499 is deleted or

FT replaced with a fragment corresponding to

FT this fragment; claim 30"

FT MISC-difference 32..38

FT /label="loop 1 modification region

FT /note="preferred region where at least one amino acid

FT residue of a parent alpha-amylase (used as a

FT template for a variant) corresponding to 14-15

FT of AAW14499 is deleted or replaced with a

FT fragment corresponding to this fragment; claim

FT 32" MISC-difference 66..84

FT /label="loop 2 modification region

FT /note="at least one amino acid residue of a parent

FT alpha-amylase (used as a template for a variant)

FT corresponding to 44-57 of AAW14499 is deleted or

FT replaced with a fragment corresponding to

FT this fragment; claim 18"

FT MISC-difference 70..78

FT /label= loop 2 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 48-51
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 20" Misc-difference 98..210
FT /label= loop 3 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 117-185 of AAW14499 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 24"
FT Misc-difference 102..206
FT /label= loop 3 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 121-181
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 26" Misc-difference 121..181
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to this fragment is deleted or
FT replaced with a fragment corresponding to
FT 102-206 of AAW14499; claim 41"
FT Misc-difference 121..174
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to this
FT fragment is deleted or replaced with a fragment
FT corresponding to 102-199 of AAW14499; claim 42"
FT Misc-difference 165..177
FT /label= loop 3 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 195-202 of AAW14499 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 21"
FT Misc-difference 166..173
FT /label= loop 3 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 196-198
FT of AAW14499 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 23" Misc-difference 181..184
FT /note= "an amino acid fragment corresponding to this
FT region is deleted from the parent sequence of a
FT variant Fungamyl; claim 43"
FT Misc-difference 291..313
FT /label= loop 8 modification region
FT /note= "at least one amino acid residue of a parent
FT alpha-amylase (used as a template for a variant)
FT corresponding to 322-346 of AAW14498 is deleted
FT or replaced with a fragment corresponding to
FT this fragment; claim 36"
FT Misc-difference 297..313
FT /label= loop 8 modification region
FT /note= "preferred region where at least one amino acid
FT residue of a parent alpha-amylase (used as a
FT template for a variant) corresponding to 325-345
FT of AAW14498 is deleted or replaced with a
FT fragment corresponding to this fragment; claim
FT 38"
XX MO9623874-A1.
XX PD 08-ATG-1996.
XX PF 05-FEB-1996; 96WO-DK00057.
XX PR 10-NOV-1995; 95DK-0001256.

PR 03-FEB-1995; 95DK-0000128.
PR 23-OCT-1995; 95DK-0001192.
XX (NOVO) NOVO-NORDISK AS.
XX Bisgard-frantzen H, Borchert TV, Svendsen A:
XX WPI: 1996-371424/37.
DR
XX
PT Alpha-amylase variants and methods of production - have altered
PT properties such as calcium dependency, substrate binding and
PT stability
XX
PS Disclosure; Page 87-88; 171pp; English.
XX
CC The present sequence is the mature Aspergillus oryzae alpha-amylase (A).
CC Variants of parent termamyl- and fungamyl-like alpha-amylases (and
CC methods of constructing them) are claimed. Examples of variants are
CC featured above. The variants have altered properties such as calcium
CC dependency, substrate binding and stability. Also one or more proline or
CC cysteine residues in the variant is modified or replaced with a
CC non-proline or non-cysteine residue such as alanine. The variants can be
CC used for (dish)washing, as detergent additives or for fabric desizing or
CC starch liquefaction. They can also be used for the production of
CC sweeteners and ethanol from starch. See also AAW14498-99.
XX
SQ Sequence 478 AA:

Query Match 95.8%; Score 2572; DB 17; Length 478;

Best Local Similarity 99.8%; Pred. No. 3; 3e-217; Mismatches 1; Indels 0; Gaps 0;

Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 ATPADMRSGSIFLPLDREARTDGTATCNTADQKRCGGTGWGIIKIDLYIOGMGTAI 80
DB 1 acpawrsgsiyflldrfartgstatacontadqcyggtwgiidkidyqgmgtai 60
QY 81 WIPVTAOLPQTAYGDAYHGWODIYSINENYGTADLKALSLAHERKATLWYVA 140
DB 61 wlpvtlaqlptlaygdahygywgdylslnenygtadlkalssalhergmylmwdiva 120
QY 141 NHMKYDAGSSVDSVSVKRPSSODYFHPFCFIONYEDQOVEDCWLGDNTVSLPDDITTK 200
DB 121 nhmydgagssvdsvsvkrrpssqdythpfcfignyeoqtwedwlgdntvslpdditk 180
QY 201 DVVKNENYDWGSLVSNYSIDGLRIDTVKHVQKDFWPGYKKAAGVYCIGEVLDGDPAYTC 260
DB 181 dvvknnewdwgslsvsnysidglridtvkhvqkdfwpgynkaagvyvcigevldgdpaytc 240
QY 261 PYQNVMDGVINYPPIYPLLAHFKSTSGSMDLNMINTVKSQCPDSTLLGTPVENHNDNR 320
DB 241 pyqnvmdgvlnyppllypllnaifkstsghmdllynmintvksdcpdstllgtfvenhndnr 300
QY 321 FASTNIDIALKANNAARFIINDGIPITVAGOEHOHYAGNDPANREATWLSGYPTDSELYK 380
DB 301 fastndialaknaaafiindgipitvagaqeyagndpanreatwlsgyptdseleyk 360
QY 381 LIASANAIRNYAISKDTGFVYKKNPPIYKDDTTIAMKKGTGDSQIYVILSNKASGDSYTT 440
DB 361 liasanaairnyaikdgtfvykknppiykddittiamkxgdsgqivylsnkagsdsytl 420
QY 441 LSLGAGYTAQOQLTEVIGCTTVVSGDNVPVPMAGCLPRVLYPTBKLAGSKTCSS 498
DB 421 lslgagyltaqqltevigcttvvsgdgvpmagclprvlypteklagskicss 478

RESULT 5
AAR46065
ID AAR46065 standard; protein: 478 AA.
XX
AC AAR46065;
XX
DT 18-JUL-1994 (first entry)


```

XX DE Mutant alpha-amylase.
XX KM Methionine substitution; stability; activity; detergent;
XX KW dishwashing agents; lification agents.
XX OS Aspergillus oryzae.
XX PN WO9402597-A.
XX PD 03-FEB-1994.
XX PF 06-JUL-1993; 93WO-DK00230.
XX PR 23-JUL-1992; 92DK-0000946.
XX PR 16-DEC-1992; 92DK-0001503.
XX PR 15-MAR-1993; 93DK-0000292.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI Bisgard-frantzen H, Svendsen A;
XX DR WPI; 1994-048855/06.
XX PT Mutant alpha-amylase from Bacillus species comprising a
XX PT methionine substitution - with improved stability and activity at
XX PT low pH, for use in detergents, dishwashing agents and
XX PT lification agents
XX PS Claim 1; Page 7; 20pp; English.
CC CC The sequence os that of the Asoergillus oryzae alpha amylase, sold
CC CC commercially as FUNGAMYL (TM) by Novo Nordisk A/S. The sequence can
CC CC be mutated by substitution of one or more of its methionine residues
CC CC for any amino acid other than cysteine. The mutant alpha-amylase
CC CC exhibits a better activity level and better stability in the
CC CC presence of oxidising agents than previous mutant alpha amylases,
CC CC and improved thermostability at moderately low pH. The enzyme can
CC CC be used as an additive for detergents, dishwashing agents and
CC CC lification agents.
SQ Sequence 478 AA;

Query Match 95.1%; Score 2552; DB 15; Length 478;
Best Local Similarity 99.4%; Pred. No. 3e-215;
Matches 475; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 21 ATPADMRQSISYFLITDRFARTDSTTATCNTADOKYCGGTWOGIIDKLDYTOGMGFTAI 80
DB 1 atpadmrqsisyflitdrfartdsgstlactadqkyggtwgilidkldyigmgftai 60
OY 81 WITPVTADLPQTAYAGDAHYGWODIYSLNENYGTADDLKALSSALHERGMYLWVDVYA 140
DB 61 witpvtadlpqtiayagdahywgqdiyslnenygtadddlalssalhergmylmvdvya 120
OY 141 NHMGTDGSSVDVSVEKFPSSQDYFHFPCEIQNYEDQYVEDCWLGDNTVSLPDLDTTK 200
DB 121 nhmgtdgssvdvsvekfpsdqdyfhfpceiqnyedqyvedcqlgdntvslpdltdtk 180
OY 201 DVVKNEWDMWGSIVNSNSTGLRIDTVKHOKDFWFGYKKAACVYTCIGEVLDGDPAYTC 260
DB 181 dvvknewdmwgsivnsnstglridtvkhvqkdfwpgynaagvycigevldgdpaytc 240
OY 261 PYONMDEVNLNPIYIYPLNFAKSTSGSMDLYNMNINTVKSDCPSTLLGFVFNHNP 320
DB 241 pyonmdevnlpiyiypllnafkstsgsmdlynmnintvscdpscllgtfvenhnpr 300
OY 321 FASTTNDIALAKNNVAFTIINDGIPITTYAGGEQHYAGCNDPANREATWLSGYPTDSELYK 380
DB 301 fastyndialaknvaafitindgipittyaggeqhyagcndpanreatwlsgyptdselyk 360
OY 381 LIASANAIRNAISKDGFTVYTKNMPYIKDDTTIAMKKGIDGSGIVITILSKGASGDIY 440

```

```

DB 361 liasanairnaiskdgftvrytknmpykddittlamkkgcdsgivtilnskasgdsyt 420
OY 441 LSLSGAGTAAQOLTEVIIGCTTWTGSDGNVPVPMAGLPRVLYPTEKLAKGSKICSS 498
DB 421 lslsgagtaaqqltevigcttvtwsgdgnvpvmagglprvlypteklsgskicss 478

RESULT 6
AAR79025
ID AAR79025 standard; protein; 478 AA.
XX
AC AAR79025;
XX
DT 22-MAR-1996 (first entry)
XX
DE Mature taka-amylase A.
XX
KW Wild type; neopullulanase; B. stearothermophilus; mutant;
KW food industry; modification; hydrophobicity; replacement
KW insertion; deletion.
XX
OS Aspergillus oryzae.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 30..38
FT Disulfide-bond 150..164
FT Disulfide-bond 240..283
FT Disulfide-bond 439..474
XX
PN JP07177891-A.
XX
PD 18-JUL-1995.
XX
PF 31-OCT-1994; 94JP-0288658.
XX
PR 12-NOV-1993; 93JP-0306096.
XX
PA (EZAK ) EZAKI GLICO CO.
XX PA (NIDE ) NEC CORP.
XX
DR WPI; 1995-279919/37.
XX
PT Modifying a transferase by enhancing hydrophobicity of a selected
PT site - increases transfer activity, also new mutant
PT neo-pullulanase(s)
XX
PS Disclosure; Page 10-11; 18pp; English.
XX
CC This sequence represents the mature form of taka-amylase from A. oryzae.
CC This sequence was used in a method for the generation of mutant
CC pullulanases for use in the food industry (see also AAR79026-28). The
CC wild type pullulanase enzyme was modified by a selected site of the
CC invention for enhancing the hydrophobicity of a selected site of the
CC pullulanase. The method comprises replacement of a group in the
CC selected site with a hydrophobic group, replacement of an amino acid
CC with a hydrophobic amino acid, and/or insertion or deletion of a
CC to produce neopullulanases T377F, S422V and M375L.
XX
SQ Sequence 478 AA;

Query Match 93.3%; Score 2503.5; DB 16; Length 478;
Best Local Similarity 97.9%; Pred. No. 5.5e-211;
Matches 468; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

OY 21 ATPADMRQSISYFLITDRFARTDSTTATCNTADOKYCGGTWOGIIDKLDYTOGMGFTAI 80
DB 1 atpadmrqsisyflitdrfartdsgstlactadqkyggtwgilidkldyigmgftai 60
OY 81 WITPVTADLPQTAYAGDAHYGWODIYSLNENYGTADDLKALSSALHERGMYLWVDVYA 140

```

```

Db      61  wltptvaqlpdcacaydaytgywtldyslnenygtaddlkalasalhergmymlvavta 120
QY      141  NHMGTIDAGSSVDYSVKFPSSQDYFHFPCFIQNYEDQTVQEDCWLGDNMYSLPDLDTTK 200
Db      121  nhmgydgagssvdyvfkpfssqdyfhpfcfiqnyedqtvqedcwlgnlvtslpdltdtk 180
QY      201  DVVKKEWVDWVGSIVSNYSIDGLRIDYVKAHQKDFWPGYNKAAGVYCAGEVLDDGPAYTC 260
Db      181  dvvkewvdwvgsivsnysidglridtkvkhqkdfwpgynkaagvytcgelvdgdpaytc 240
QY      261  PYQNVMDVLYNPIYYPPLNAFKSTSGSMDLYNMINTVKSDCPDSTLLGTFVENHNP 320
Db      241  pyqnmvdgvllypyypllnafkstsgsmdllynmintvksdcpdstllgtfvenhnp 300
QY      321  FASVNDIATLAKNVAFTIINDGPIITYAGQOHYAGNDPANREATWLSGPPDSLEYK 380
Db      301  fasyndiatlaknvaftiindgpiityaqeghyagndpanreatwlsqppdselyk 360
QY      381  LIASANAIRNVAISKDTGFVTYKMWPIYKDDTTIAMRKGTGSOIVTILSNKASGDSYT 440
Db      361  liasanaairnvaiskdtgfvtcykn-pykddttiamrtgtdgsqvtllsnkgsagdsyt 419
QY      441  ILSGAGYTAGQQLTEVIGCTTWTVSGDNVPVPMAGGLPRVLYPTEKLAGSKICSS 498
Db      420  lslsagysytagqltevigctvtvsgdnvpmagglprvlypteklagskicsds 477

```

```

RESULT 7
AAR88212
ID      AAR88212 standard; Protein: 493 AA.
XX
AC      AAR88212;
XX
DT      03-APR-1996 (first entry)
XX
DE      Alpha-amylase.
XX
KW      Alpha-amylase; thermostable enzyme; baking; Aspergillus niger.
OS      Thermomyces lanuginosus CBS 224.63.
XX
FH      Key
FT      Location/Qualifiers
ET      1..18
PT      Peptide
XX
XX      W09601323-A1.
XX      18-JAN-1996.
XX      03-JUL-1995; 95WO-EP02607.
XX      04-JUL-1994; 94GB-0013419.
XX      (DANI-) DANISCO AS.
XX      Michelsen B, Rasmussen P;
XX      MPI. 1996-087673/09.
XX      N-PSDB; AAT10562.
XX
XX      Thermophilic alpha-amylase with activity range of 60-80 degrees C -
PT      derived from Thermomyces lanuginosus, useful in the prepn. of
PT      foodstuffs and bakery prods. esp. bread
XX
PS      Claim 3: Page 36-38; 94pp; English.
XX
XX      A thermostable alpha-amylase (AAR88212) of Thermomyces lanuginosus
CC      CBS 224.63 is expressed in Aspergillus niger hosts using a gene
CC      (AAT10562) isolated from a T. lanuginosus gene library. The
CC      recombinant enzyme (54-60 kDa) shows optimal activity at
CC      60-70 deg and pH 5.8-6, has a pI of 3.7 and is active at 60-80
CC      deg.
XX
XX

```

SQ Sequence 493 AA:

Query Match 62.1%; Score 1666.5; DB 17; Length 493;
 Best Local Similarity 63.2%; Pred. No. 17e-137;
 Matches 308; Conservative 69; Mismatches 101; Indels 9; Gaps 4;

```

QY      16  APAL--AAPADRRSOSIYFLTDRFAPRDGSTTATCNTADKRCGCGTWQGIIDRLDYIQ 73
Db      12  spllyraaapdekaqslfmltdrfardcnstapcdtaqkyqgvtgrjlinldyiq 71
QY      74  GMGFTAIWTPVTAQLPQTTAVGDA--YHGWOODIYSINENYGTADLKAISALHERG 131
Db      72  dmgtfaiwtpvtaqwdvdaadaatsyngyqkdlyslnskfytaddlkaladaahary 131
QY      132  MYLMDVYVNMHMGYDAGSSVDYSVKFPSSQDYFHFPCFIQNYEDQTVQEDCWLGDNMY 191
Db      132  mlmdvvanhlgysgshsevdyslfnplnsqdyfhpfcfiiedydngevegcwladpct 191
QY      192  SLPDLPTKDVKNENYDWVGSIVSNYSIDGLRIDYVKAHQKDFWPGYNKAAGVYCIGEV 251
Db      192  tlpdvtltpqvvtffndwkslvanysidglrvctkvhekdfwpdfneaac-tvgev 250
QY      252  LDGDPAYTCFYQNVMDGVLYNPIYYPPLNAFKSTSGSMDLYNMINTVKSDDPDSTLLGT 311
Db      251  lngdpaytcpyqvgvldgvlhpyypaldafksgvnglglagaltvgesckdnllygn 310
QY      312  FVENHNPRASTYNDIATLAKNVAFTIINDGPIITYAGQOHYAGNDPANREATWLSG 371
Db      311  flenhdiafasyddiatlaknglatilisdgipilytqeghyagndpanreawls9 370
QY      372  YPTDSELYKLIASANAIRNVAISKDTGFVTYKMWPIYKDDTTIAMRKGTGSOIVTILSN 431
Db      371  yntdaelygfikangirnaisqneptsckkvyqgdstclainrg---gvvtvln 426
QY      432  KGASGDSYFSLSGAGYTAGQQLTEVIGCTTWTVSGDNVPVPMAGGLPRVLYPTEKLAG 491
Db      427  egaagelgtvslpigtfeagteltdvlscktlvagsagadvlpsslpvlypsqjak 486
QY      492  SKICSS 498
Db      487  sgicasa 493

```

```

RESULT 8
AAR07574
ID      AAR07574 standard; protein: 512 AA.
XX
AC      AAR07574;
XX
DT      02-FEB-1991 (first entry)
XX
DE      Alpha-amylase encoded by the AMY1 gene.
XX
XX      Schwanniomycetes yeast cells; expression cassette; alpha-amylase;
KW      marker gene; region; signal peptide; terminator;
KW      autonomously replicating sequence.
XX
XX      Schwanniomycetes occidentalis.
OS
XX
XX      Key
FH      Location/Qualifiers
FT      Disulfide-bond 66..74
FT      /note="1"
FT      Disulfide-bond 186..200
FT      /note="2"
FT      Disulfide-bond 276..319
FT      /note="3"
FT      Disulfide-bond 475..510
FT      /note="4"
XX
XX      EP394538-A.
XX      31-OCT-1990.
XX
XX

```

XX	28-APR-1989;	89EP-0107780.
PF		
XX	28-APR-1989,	89EP-0107780.
PR		
XX		
PA	(RHEI) RHEIN BIOTECH GFS.	
XX		
PI	Hollenberg C, Strasser A;	
XX		
DR	WPI; 1990-328670/44.	
DR	N-PSDB; AA006388.	
XX		
PT	Transformed Schwannlomyces yeast cells - contg. an expression	
PT	cassette contg. regulon, DNA coding for foreign protein and	
PT	terminator	
XX		
PS	Disclosure; Fig 1B(1-6); 59pp; English.	
XX		
CC	The expression cassette may contain a regulon (R), a signal peptide	
CC	sequence (S), a foreign protein sequence (F) and a terminator (T)	
CC	of which R and/or S and/or T are derived from the alpha-amylase gene	
CC	of S. occidentalis.	
CC	R is pref. a 1.8 kb BglII-XhoI fragment (bases -1 to -540 of this	
CC	sequence).	
CC	S contains all or part of one of the sequences represented	
CC	in AA006383-87.	
CC	F is e.g. a cellulase, interleukin, insulin-like-growth factor,	
CC	interferon etc.	
CC	T is pref. all or part of the terminator of this sequence	
CC	(bp 1537-1740).	
CC	AW1 acts as selective marker for the vector carrying the expression	
CC	cassette.	
CC	See also AA006389.	
XX		
XX		
XX	Sequence 512 AA;	
XX		

Query Match	54.0%;	Score 1449;	DB 11;	Length 512;
Best Local Similarity	56.4%;	Pred. No. 2.3e-118;		
Matches 269;	Conservative 74;	Mismatches 132;	Indels 2;	Gaps 2

OY	21	ATPAD-WRSOSTYFELITDRFAPRTGOSTTATTCNADOKYCGGTMOGILDKLDYIOGMGFTA	79
Db	36	ssadckwkdqslayqivltarfaidsqtsadcladysdrkycgsykyiklkladyigsmflla	95
OY	80	ITWTPYTAQIPGPTVAGAYAGWYOOOYLSINENYEGMADLALSSALHEGMYLMDVY	139
Db	96	twspvveqipontaygyayayngymkntidencntnftgaeklqaselhstrsmllmvdv	155
OY	140	ANHKGTGAGSSVDYSVFEKPESSODYFHPPECFTQNTEDQTYQEDCWLGDNTVSLPDLDTT	199
Db	156	ynhyamngdssvdyssfcfptngsqsyfhydcylltnydcqlunvdcwegdewtsipdlste	215
OY	200	KDYYKNEWDWNGSLYSNNTSIDGLRIPYHNYOCDKWPNGNKAAGYCYCIGFVLDGDPAYT	259
Db	216	dneiyvigtvwsdfvqngsytoglrldakrkvdtasltkfedaasygmnygeyqddpyt	275
OY	260	CPYONWMDGYLNPPIYPLPLAFAKSTGSGSMDDLYNMINTVYSDPCDPSLTGLTFVENHDNP	319
Db	276	cpyygmkytynpipyfifsdatsatseilmistlqscsdvalllgnfienhdqv	335
OY	320	RFASYNMIDALAKNNAAPFIINDGIPITTYAGOBQDHAAGDANRANRATWLSYPIPDSELY	379
Db	336	rfspsvtsdtsllhndmaflilldgdpilpyllyyqeqglingdsdpantreaawlsgynidseyy	395
OY	380	KLASANAIENYNAISKDTEFYVYKWPPIYKDDTITLARKGTGOSQVITLTSNKGSGSY	439
Db	396	ellsklnglntqkikdaesystrkssvssedhyatrsgsantqljstfmlngnsq	454
OY	440	TLSISGAGTGAQQLTEVYIGCTTYTVSGDNVYPPAAGCIPRYLYTEKLAGSKTCS	496
Db	455	dlvsnctvgyssgdkvdlvllscsvtlaagsgslvslsggmppgyavssvslsggcn	511

RESULT	9
AAP81180	
ID	AAP81180 standard; protein; 511 AA

DT 12-DEC-1990 (first entry)

DE Sequence of alpha-amylase encoded by genomic DNA of Schwanniomycetes
DE castellii.

KW Brewing; beer; breadmaking; biomass.

Schwanniomycetes castelii.

PN EP260404-A.

PD 23-MAR-1988.

PF 17-JUL-1987; 87EP-0110370.

PR 17-JUL-1987; 87EP-0110370.

PA (HETB) HEINEKEN TEC BEHEER NV.

PI Strasser A, Martens FB, Dohmen J, Hollenberg CP;

DR WPI; 1988-078794/12.

XX

PT esp. coding alpha amylase or glucoamylase, able to convert

XX 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040

PS Claim 18; Flg 2A-2C; 74pp; English.

CC A cosmid library is constructed from genomic DNA of ATCC26076 and this
CC screened for inserts confg. the *AG* gene by ability to transform
CC *S.cerevisiae* GRI 18 to histidine prototrophy. One positive cosmid
CC (32kb: pvc1) was digested with *EcoRI*, religated and used to transform
CC *E.coli* JRA21. Positive transformants contain pvc1- α , which includes
CC a 5kb *EcoRI* fragment confg. the gene in AA081525.

Sequence 511 AA;

Query Match	52.2%;	Score 1400.5;	DB 9;	Length 511;
Best Local Similarity	55.1%;	Pred. No. 4.1e-114;		
Matches 264;	Conservative 74;	Mismatches 134;	Indels 7;	Gaps 4;

```

Oy 21 A1PAD-WRSOSIYELLTDREARFDGSTTANCTADOKYCGGTWOGIIDLKDIYQMGFTa 79
    ::|||::||||::|||||::|:||||:::|||||||||
Db 36 sssadkkdkdgslyqivtarfarsdgttdclvsdrtkycggsykgjldkdyiqgmfta 95

```

QY 80 IWTPVTAQLPQTAYGDAYHGYWQDIYSINENYGTADDIKALSSALHERGMYLMD-- 137

Db 96 i wispvveqipdntaygyayhgywmknidelntnfgtadelkqglaselhfrsmymy---dgr 152

QY 138 VVANHMGYDGAGSSVDYSEVKPFSSQDYFHPFCFIQNYEDQTQVEDCWLGDNTVSLPDL 197

Db 153 rcynhyawngdssvdysftcfnqgsyfhndyclitnyndqtnvedcwegdtevs1pdlis 212

QY 198 TTKDVKNEMYDWGSLVSNYSIDGLRIDTVKHVQKDFWPBGYNKAAGVYCIGEVLDGDP 257

Db 213 tednevigftwsdfvqnsidqlrldsakhvdtasltkfedasgvynlgevygddpt 272

QY 258 YTCPYQNVMDGVLNYPRIYPLLNAFKSTSGSMDDLNMINTVKSDCPDSTLLGTFEVNHHD 317

Db 273 ytcpqnymkvtnpplypvyrffsdsatsseltsmistlqsscdrvslqnfnhd 332

QY 318 NPREASYTNDIALAKNVAFIILNDGIPITYAGQEOHYAGGNDPANREATWLSGYPTDSE 377


```

PR 04-MAR-1993; 93JP-0069303.
XX
XX (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
PA
XX
XX WPI: 1994-328987/41.
DR N-PSDB; AA077665.
XX
XX Variant carbohydrate hydrolase(s) with increased activity -
PT consists of e.g. alpha-amylase with tyrosine residue in enzyme
PT centre, useful for mass-prodn. of oligosaccharide(s)
XX
XX
PS Example 1; Page 13-15; 27pp; Japanese.
XX
XX AAR63184-7 show the amino acid sequences of variant alpha amylases,
CC composed by substituting bases 247-249 of the structural gene (see
CC AA077665-8) region, with TTC, TGC, CTC or AAC. These substitutions
CC result in the 83rd amino acid residue (tyrosine) of the wild type
CC sequence being changed to phenylalanine, tryptophan, leucine or
CC asparagine respectively. The substituted amino acid is present in
CC the active site of the enzyme and confers increased activity on the
CC enzyme. The variants are useful for the mass production of
CC oligosaccharides. (see AAR63188 for the variant structure of a
CC cyclomaltohexarin glucanotransferase).
XX
XX
SO Sequence 468 AA:

Query Match 51.9%; Score 1392; DB 15; Length 468;
Best Local Similarity 54.4%; Pred. No. 2e-113;
Matches 259; Conservative 81; Mismatches 126; Indels 10; Gaps 5;

QY 22 TPAD-WRSOSIYFLTLDRFARTDGTATCNTADKCYCGGTGQGIIDKLDYIGMGFTAI 80
DB 2 tndakwrsqsiyivdrfartdgtatcncntadkcycggtgqgiikldyikmgftai 61
QY 81 WTPVTAQLPQTTAYGDAYHGWOODIYSLNENYGTADDLKALSALHERGMYLMDVVA 140
DB 62 wisprvenipdntatayayahgwmknikylnenfgtaddklsaqelhdrcdmllmwdivt 121
QY 141 NHMGYDGAASSVDYVYKRFSSODYFHPFCFIONTEDOTQYEDCWLGDNTVSLPDLDTK 200
DB 122 nhysdgsqdsidyseyrpfnndqkyfhnyclisnyddqagvqscwegssvalpdlrted 181
QY 201 DAVKKNEMYWGSVLNYSNIDGLRIDTQKHKVOKDFWPGCNKAGVYICGEVLDGDPAYTC 260
DB 182 sdvaasvfnswvdiqgnysidglridtsakhvqdgffpdrfvaasvgyvgefgqdpaytc 241
QY 261 PYONVMDGYLVNPIYPLINAFKSTSGSMDLLYNNIMTKSDCSPSTLLGTFVENHNDPR 320
DB 242 pyqnyipgvsynpilyypttrffctdssaselmqmssvaascsdprclltufvenhndpr 301
QY 321 FASYTNDIALAKNVAFIILNDGIPITAYAGOHYAGGNDPANREATYWSGYPDSLEYK 380
DB 302 fasmtdqslisnataiatyilgdiqipvlyygeqglsqskpnmrealtwlsygnksedyk 361
QY 381 LIASANAIRNVAISKDTGFVTYKNNPIYKDDPTIAMRKGTGDSOIIVTILSKGASGDSYT 440
DB 362 likamtsdqsilsnataiatyilgdiqipvlyygeqglsqskpnmrealtwlsygnksedyk 361
QY 441 LLSGAGYTAGOOLTEVICCTTVYVGSDDGNVPVPAAGLPRVLYPTEKSLASGDSYT 496
DB 417 vtlnstgysgedlvevlfcstlvsygsd--lqvslqggqqlfypak--ysadlcs 468

RESULT 14
AAR63185
ID AAR63185 standard; Protein; 468 AA.
XX
XX AAR63185;
AC
XX 16-JUN-1995 (first entry)
DT
XX Variant alpha amylase deriv. from Saccharomyopsis fibuligera.
DE

```

```

XX alpha amylase; carbohydrate hydrolase; increased activity;
XX tyrosine residue; enzyme centre; mass production; oligosaccharide;
XX variant; cyclomaltohexarin glucanotransferase.
XX
XX Saccharomyopsis fibuligera.
XX
XX Key Location/Qualifiers
XX Misc-difference 83
XX FT /Label= Y83W
XX
XX JP06253836-A.
XX
XX 13-SEP-1994.
XX
XX 04-MAR-1993; 93JP-0069303.
XX
XX 04-MAR-1993; 93JP-0069303.
XX
XX (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
XX
XX WPI: 1994-328987/41.
XX DR N-PSDB; AA077666.
XX
XX Variant carbohydrate hydrolase(s) with increased activity -
XX consists of e.g. alpha-amylase with tyrosine residue in enzyme
XX centre, useful for mass-prodn. of oligosaccharide(s)
XX
XX
XX Example 1; Page 15-17; 27pp; Japanese.
XX
XX AAR63184-7 show the amino acid sequences of variant alpha amylases,
CC composed by substituting bases 247-249 of the structural gene (see
CC AA077665-8) region, with TTC, TGC, CTC or AAC. These substitutions
CC result in the 83rd amino acid residue (tyrosine) of the wild type
CC sequence being changed to phenylalanine, tryptophan, leucine or
CC asparagine respectively. The substituted amino acid is present in
CC the active site of the enzyme and confers increased activity on the
CC enzyme. The variants are useful for the mass production of
CC oligosaccharides. (see AAR63188 for the variant structure of a
CC cyclomaltohexarin glucanotransferase).
XX
XX
SO Sequence 468 AA:

Query Match 51.8%; Score 1391; DB 15; Length 468;
Best Local Similarity 54.4%; Pred. No. 2.5e-113;
Matches 259; Conservative 81; Mismatches 126; Indels 10; Gaps 5;

QY 22 TPAD-WRSOSIYFLTLDRFARTDGTATCNTADKCYCGGTGQGIIDKLDYIGMGFTAI 80
DB 2 tndakwrsqsiyivdrfartdgtatcncntadkcycggtgqgiikldyikmgftai 61
QY 81 WTPVTAQLPQTTAYGDAYHGWOODIYSLNENYGTADDLKALSALHERGMYLMDVVA 140
DB 62 wisprvenipdntatayayahgwmknikylnenfgtaddklsaqelhdrcdmllmwdivt 121
QY 141 NHMGYDGAASSVDYVYKRFSSODYFHPFCFIONTEDOTQYEDCWLGDNTVSLPDLDTK 200
DB 122 nhysdgsqdsidyseyrpfnndqkyfhnyclisnyddqagvqscwegssvalpdlrted 181
QY 201 DAVKKNEMYWGSVLNYSNIDGLRIDTQKHKVOKDFWPGCNKAGVYICGEVLDGDPAYTC 260
DB 182 sdvaasvfnswvdiqgnysidglridtsakhvqdgffpdrfvaasvgyvgefgqdpaytc 241
QY 261 PYONVMDGYLVNPIYPLINAFKSTSGSMDLLYNNIMTKSDCSPSTLLGTFVENHNDPR 320
DB 242 pyqnyipgvsynpilyypttrffctdssaselmqmssvaascsdprclltufvenhndpr 301
QY 321 FASYTNDIALAKNVAFIILNDGIPITAYAGOHYAGGNDPANREATYWSGYPDSLEYK 380
DB 302 fasmtdqslisnataiatyilgdiqipvlyygeqglsqskpnmrealtwlsygnksedyk 361
QY 381 LIASANAIRNVAISKDTGFVTYKNNPIYKDDPTIAMRKGTGDSOIIVTILSKGASGDSYT 440

```

```

Db 362 LIAANAARNAAYQDSSYATSLVIFNDHVIATKRGs---VSVFNLGSSGSS-d 416
QY 441 LSLGAGYTAAGQQLTEVIGCTTAVGSDGNVPYPMAGLPRVLVPTREKLAGSKICS 496
Db 417 VLIANTGYSAGEDLVEVLTCSTVSGSSD--LQVSIQGGPQLIVPAK--YASDICS 468

```

RESULT 15

AAR63187 standard; protein; 468 AA.

AAR63187:

16-JUN-1995 (first entry)

Variant alpha amylase deriv. from Saccharomyopsis fibuligera.

alpha amylase: carbohydrtase hydrolase; increased activity;

tyrosine residue; enzyme centre; mass production; oligosaccharide;

variant; cyclomaltoextrin glucanotransferase.

Saccharomycopsis fibuligera.

Key Location/Qualifiers

Misc-difference 83

JP06253836-A.

13-SEP-1994.

04-MAR-1993: 93JP-0069303.

04-MAR-1993: 93JP-0069303.

(AGEN) AGENCY OF IND SCI & TECHNOLOGY.

WP1; 1994-328987/41.

N-PSDB; AA077668.

Variant carbohydrtase hydrolase(s) with increased activity -

consists of e.g. alpha-amylase with tyrosine residue in enzyme

centre; useful for mass-prodn. of oligosaccharide(s)

Example 1; Page 20-23; 27pp; Japanese.

AAR63184-7 show the amino acid sequences of variant alpha amylases,

composed by substituting bases 247-249 of the structural gene (see

AA077665-8) region, with TTC, TGG, CTC or AAC. These substitutions

result in the 83rd amino acid residue (tyrosine) of the wild type

sequence being changed to phenylalanine, tryptophan, leucine or

asparagine respectively. The substituted amino acid is present in

the active site of the enzyme and confers increased activity on the

enzyme. The variants are useful for the mass production of

oligosaccharides. (see AAR63188 for the variant structure of a

cyclomaltoextrin glucanotransferase).

Sequence 468 AA;

Query Match 51.7%; Score 1387; DB 15; Length 468;
Best Local Similarity 54.4%; Pred. No. 5.5e-113;
Matches 259; Conservative 80; Mismatches 127; Indels 10; Gaps 5;

22 TPAD-WRSOSTYFLTDPRFATDGTATCNTADQKCGTMOGIIDKLDYIGMGFTAI 80

2 Lnadkwszsqslvqldfardgdtasascntedrllycgsgfqlkldylkdmgftai 61

81 WITPYTAQLPQTAAGYAGHGYGWOODIYSLNENGTADLKALSALEHGRKYLAWDYVA 140
62 wispyventipentatayayaybgnmknlyklnenfgtaddlkslqelhdtdmlmvdvlt 121

```

QY 141 NHMGYDAGSSVDYSVFKPPSSODTFHPFCFIONEDQTVEDCWLGDNTVSLPDLDTK 200
Db 122 nhysdsgsgsidyseylfndqkyfhnycilsnysdgaqvgscwegdssvalpdlrted 181
QY 201 DVVKNENYDWNVGSLSVNSYSIDGLRIDTVKHKVOKDFPGYNKAAGVYCIGEVLDGDPAYTC 260
Db 182 sdvasvfnswvkdafvngysidglrldsakhvdqgffpdrfsasgyysgevfgqdpaytc 241
QY 261 PYONVMDGVNLNPIYPLLNAPKSTGSMDDLYNMTNTVKSDDCPDSTLLGTFFVENHNDPR 320
Db 242 pygnvlypvsnypllypdltrfiktassselqgmssvasscsdpcllntfvenhndpr 301
QY 321 FASYTNDIALKKNVAARFIILNDGIPRIYAGOEQHVAGNDPANREAPTWISGVPDSELYK 380
Db 302 fasmtdsgslismaiafvlgdipwlyyqegqlsgksdpnrtrealwlsynkesdyk 361
QY 381 LIASANAIRNVAISKDGFVTYKNMPYKDDTTIARKGTDGSGQIVTILSNKASGDSYT 440
Db 362 LIAKANAARNAAYQDSSYATSLVIFNDHVIATKRGs---VSVFNLGSSGSS-d 416
QY 441 LSLGAGYTAAGQQLTEVIGCTTAVGSDGNVPYPMAGLPRVLVPTREKLAGSKICS 496
Db 417 VLIANTGYSAGEDLVEVLTCSTVSGSSD--LQVSIQGGPQLIVPAK--YASDICS 468

```

Search completed: June 13, 2002, 08:53:25
Job time: 275 sec
